Principles of Empirical Analysis

Lecture 9: Difference-in-differences

Spring 2021
Tuukka Saarimaa

Outline

- **Basic idea of difference-in-difference (DID) designs**
  - DID with two groups and two time periods
  - More general case with many time periods
- **Applications**
  - Card & Krueger (1994): classic paper on minimum wage
  - Saarimaa & Tukiainen (2015): common pool
  - Currie & Reed (2011): pollution and health
  - Other examples

DID

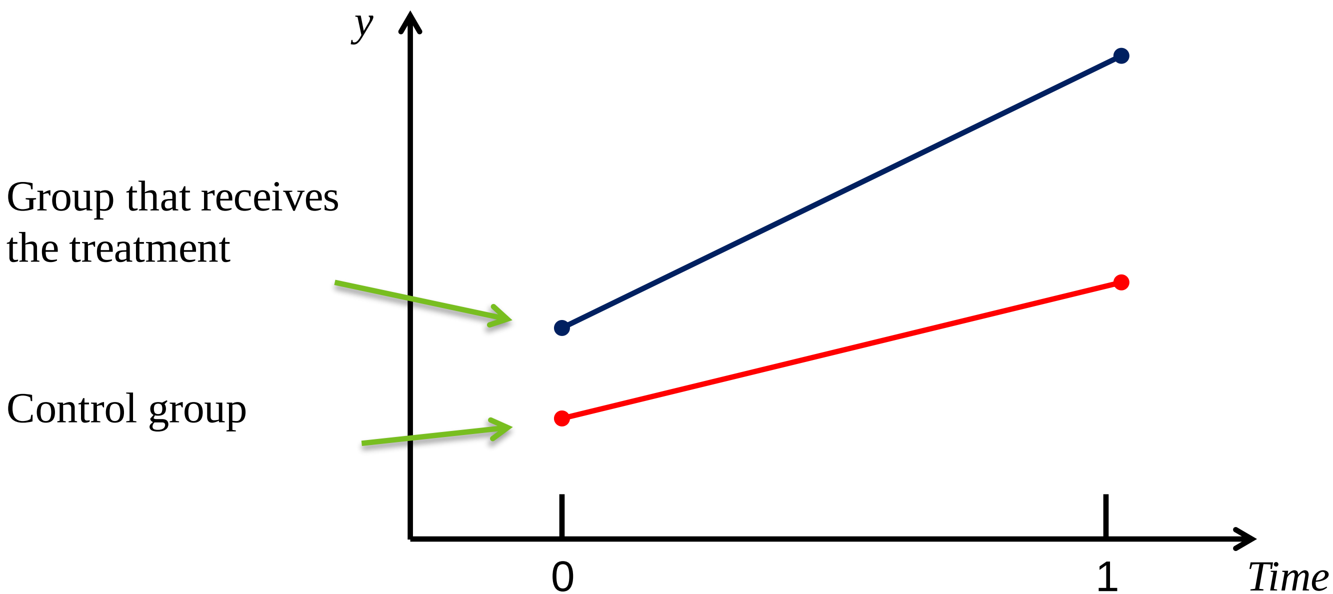
- **We have talked about the idea of using group differences to estimate causal effects**
  - We would like to find treatment and control groups who can be assumed to be similar in every way except receipt of treatment
  - Without randomization this is very difficult/implausible
- **A weaker assumption is that in the absence of treatment, the difference between treatment and control groups is constant over time (parallel or common trends)**
- **With this assumption we can use observations on treatment and control groups before and after the treatment to estimate a causal effect**

DID

- **Idea:**
    - Difference pre-treatment is ‘normal’ difference
    - Difference post-treatment is ‘normal’ difference + causal effect of treatment
    - Difference-in-differences is the causal effect
- **DID relies heavily on parallel time trends so visual inspection of the data is a very important part of any DID analysis**

Difference-in-differences – two groups, two time periods

Difference-in-differences graphically

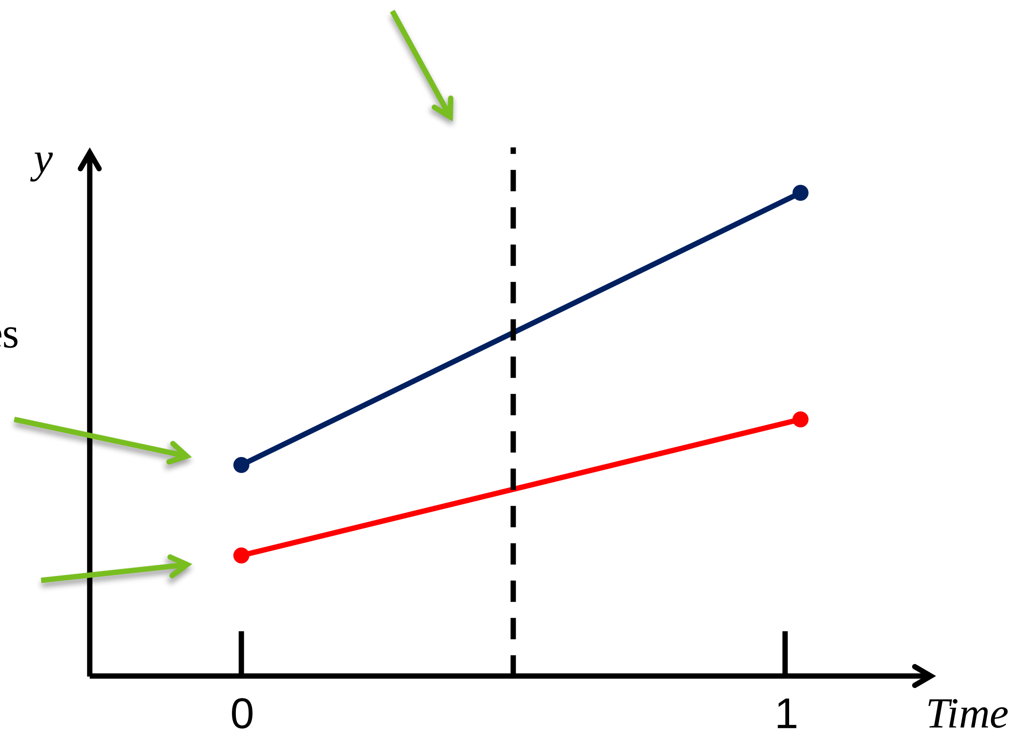


Difference-in-differences graphically

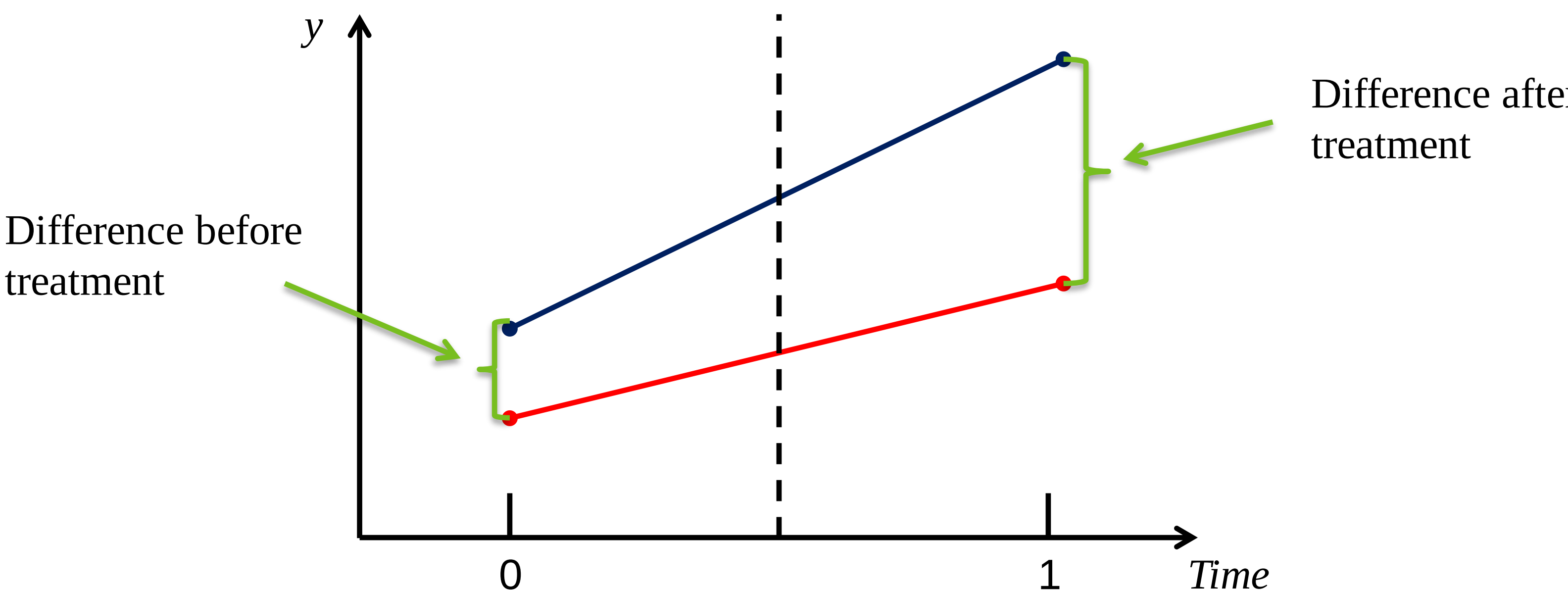
Treatment occurs but only
one group is treated

Group that receives
the treatment

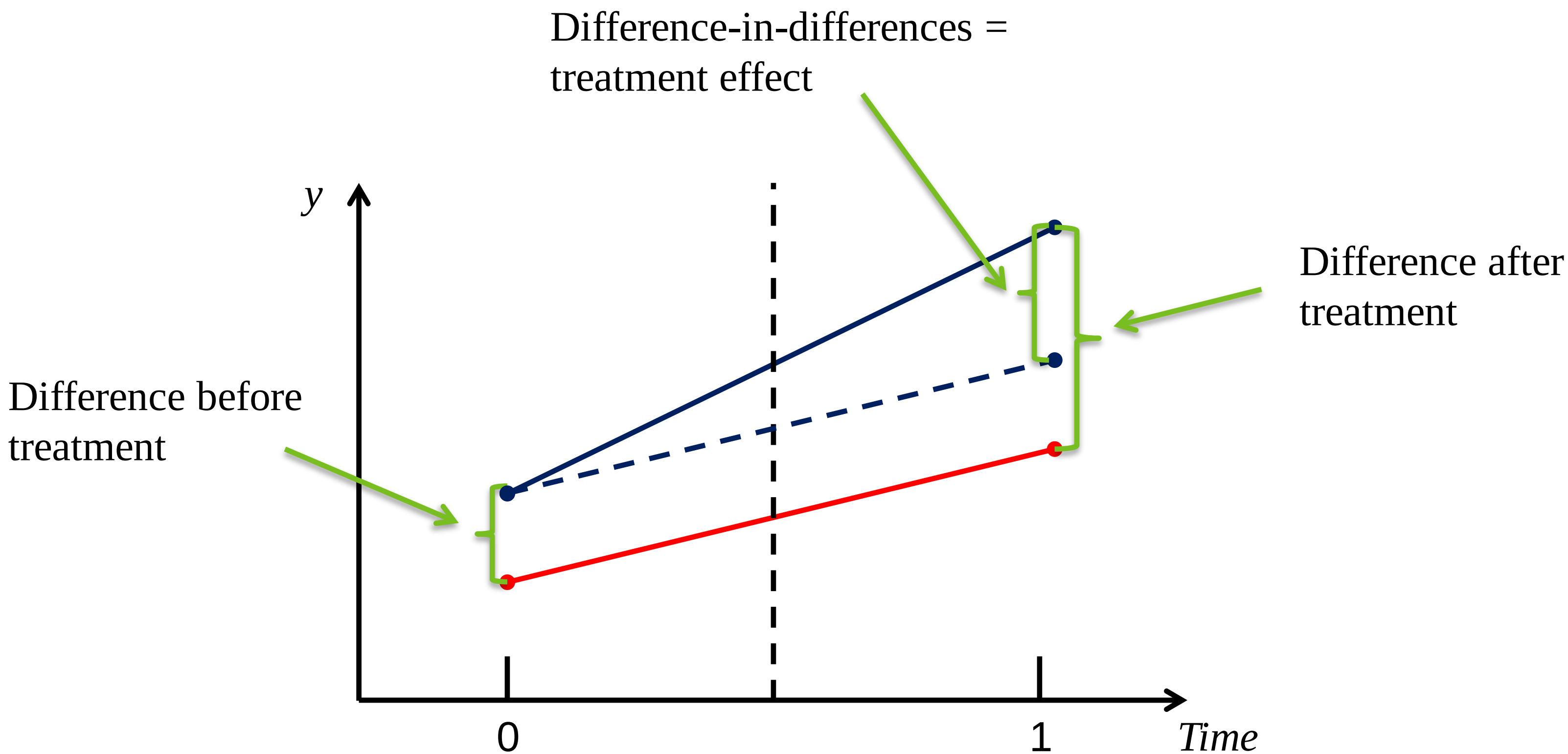
Control group



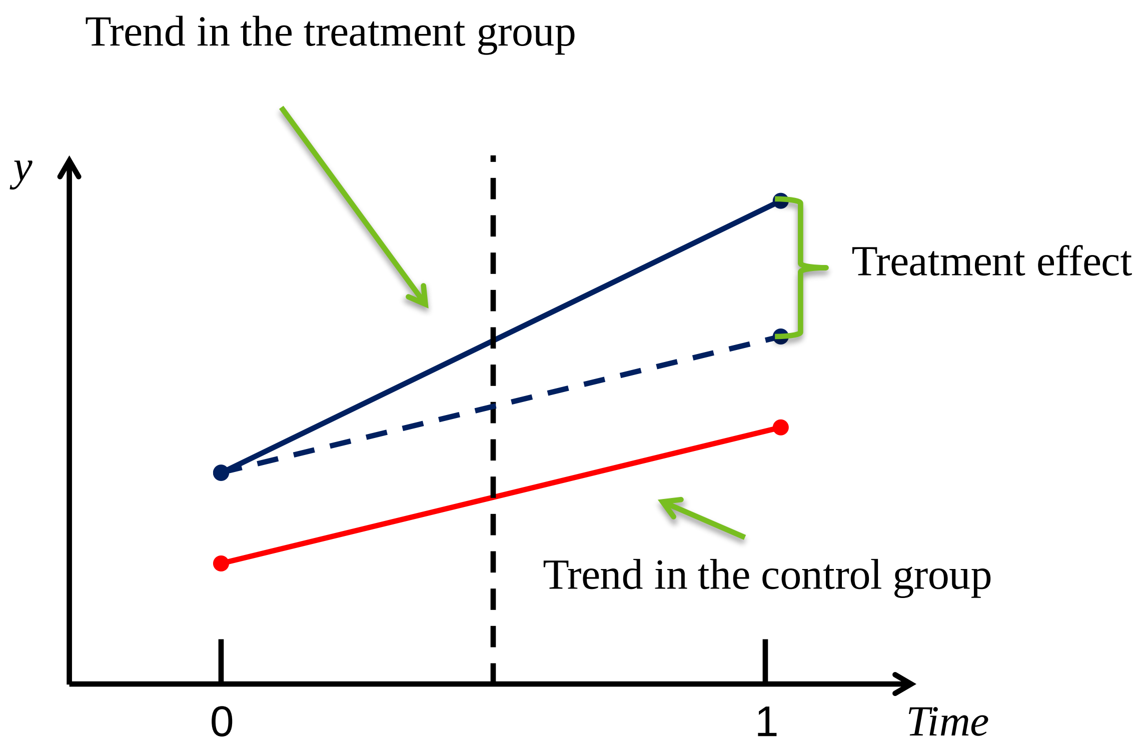
Difference-in-differences graphically



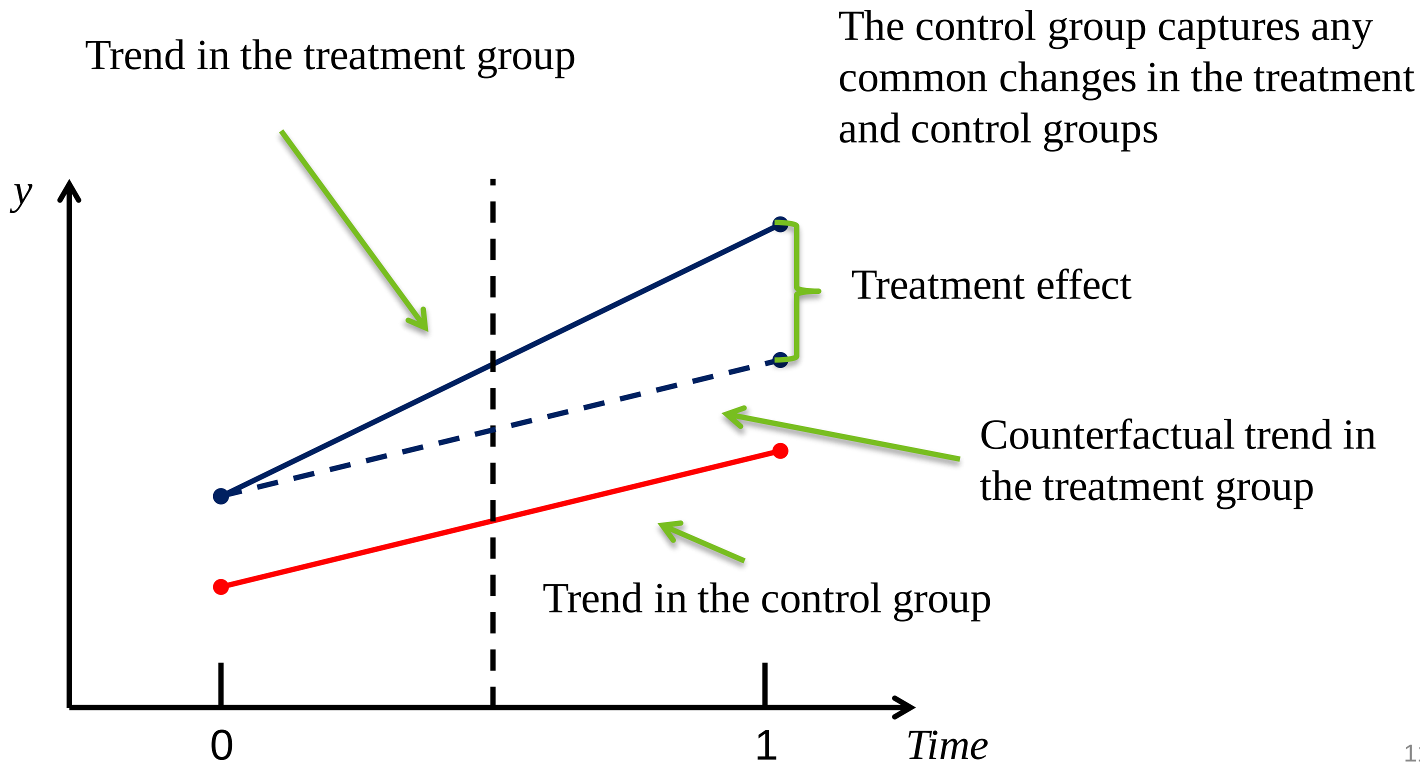
Difference-in-differences graphically



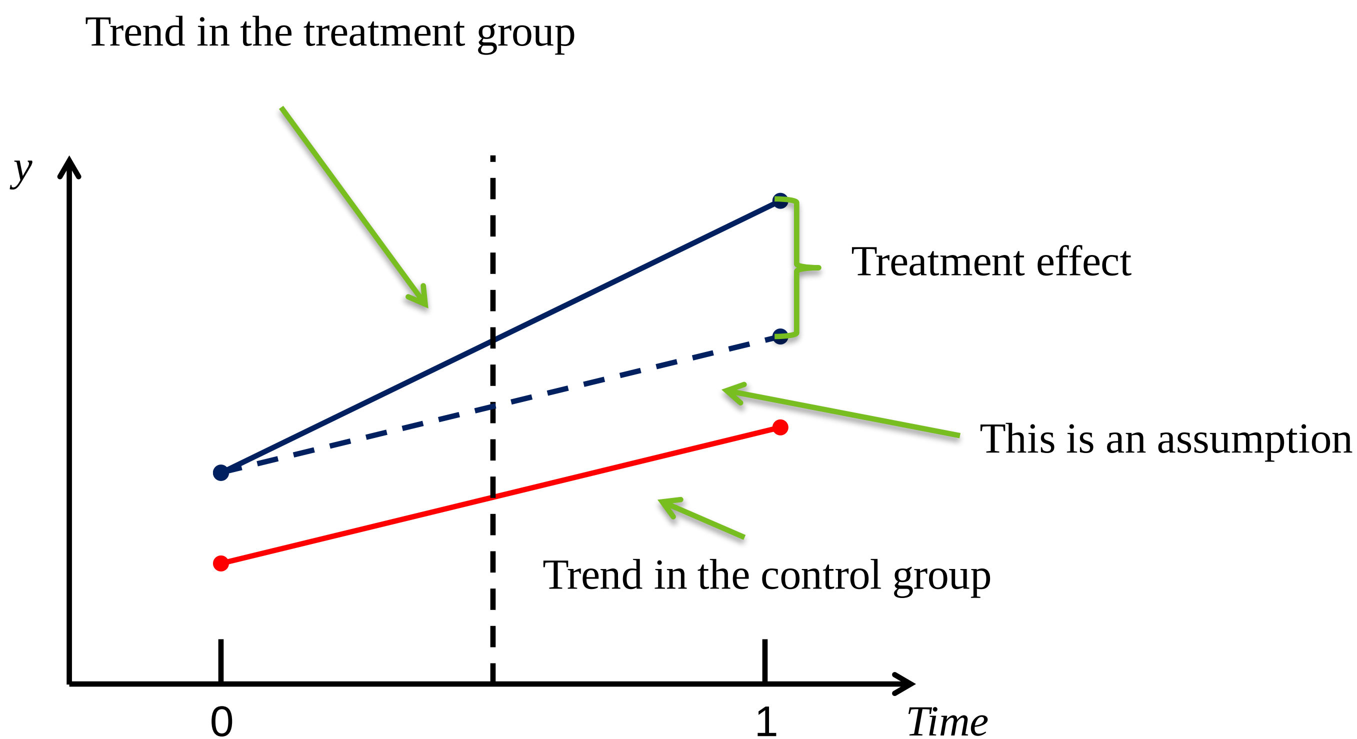
Difference-in-differences graphically



Difference-in-differences graphically



Difference-in-differences graphically



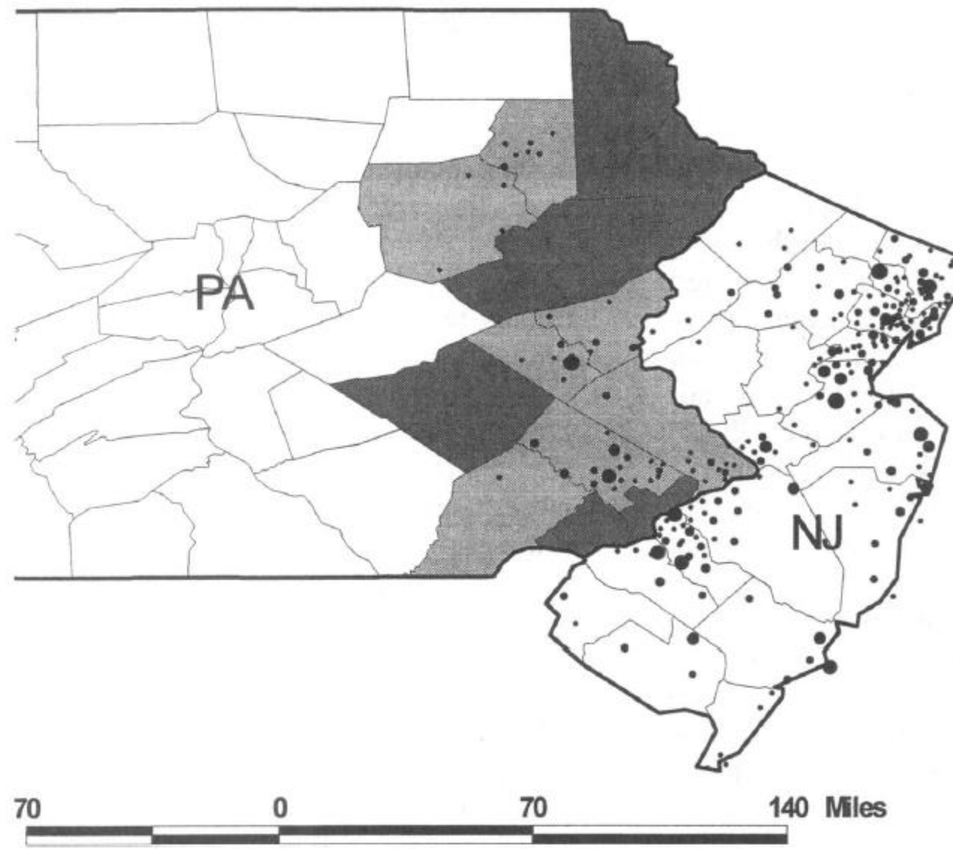
Example: New Jersey minimum wage increase

On April 1, 1992, NJ increased the state minimum wage from $4.25 to $5.05; PA's minimum wage stayed at $4.25

Card & Krueger (1994) surveyed about 400 fast food stores both in NJ and in PA before (February) and after (November) the minimum wage increase

A DID setup

Any common macroeconomic trends captured by using the control group



Card & Krueger (1994)

- **DID more formally**

  y_{ist}: employment at restaurant i, state s, time t

- **In DID we compare the following means**

  $E[y_{ist}|s = NJ, t = Feb]$

  $E[y_{ist}|s = NJ, t = Nov]$

  $E[y_{ist}|s = PA, t = Feb]$

  $E[y_{ist}|s = PA, t = Nov]$

Card & Krueger (1994)

- **In Jersey:**

$E[y_{ist}|s = NJ, t = Feb]$ = mean employment in February

$E[y_{ist}|s = NJ, t = Nov]$ = mean employment in November

$E[y_{ist}|s = NJ, t = Nov] - E[y_{ist}|s = NJ, t = Feb]$

= difference in employment

- **In Pennsylvania:**

$E[y_{ist}|s = PA, t = Feb]$ = mean employment in February

$E[y_{ist}|s = PA, t = Nov]$ = mean employment in November

$E[y_{ist}|s = PA, t = Nov] - E[y_{ist}|s = PA, t = Feb]$

= difference in employment

Card & Krueger (1994)

- **The population DID is the treatment effect we are after**

$$\delta = (E[y_{ist}|s = NJ, t = Nov] - E[y_{ist}|s = NJ, t = Feb])$$
$$- (E[y_{ist}|s = PA, t = Nov] - E[y_{ist}|s = PA, t = Feb])$$

- **The DID estimator is the sample analog:**

$$\hat{\delta} = \left(\bar{y}_{NJ,Nov} - \bar{y}_{NJ,Feb}\right) - \left(\bar{y}_{PA,Nov} - \bar{y}_{PA,Feb}\right)$$

Card & Krueger (1994)

Variable	Stores by state: PA (i)	NJ (ii)	Difference, NJ – PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	−2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	−0.14 (1.07)
3. Change in mean FTE employment	−2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

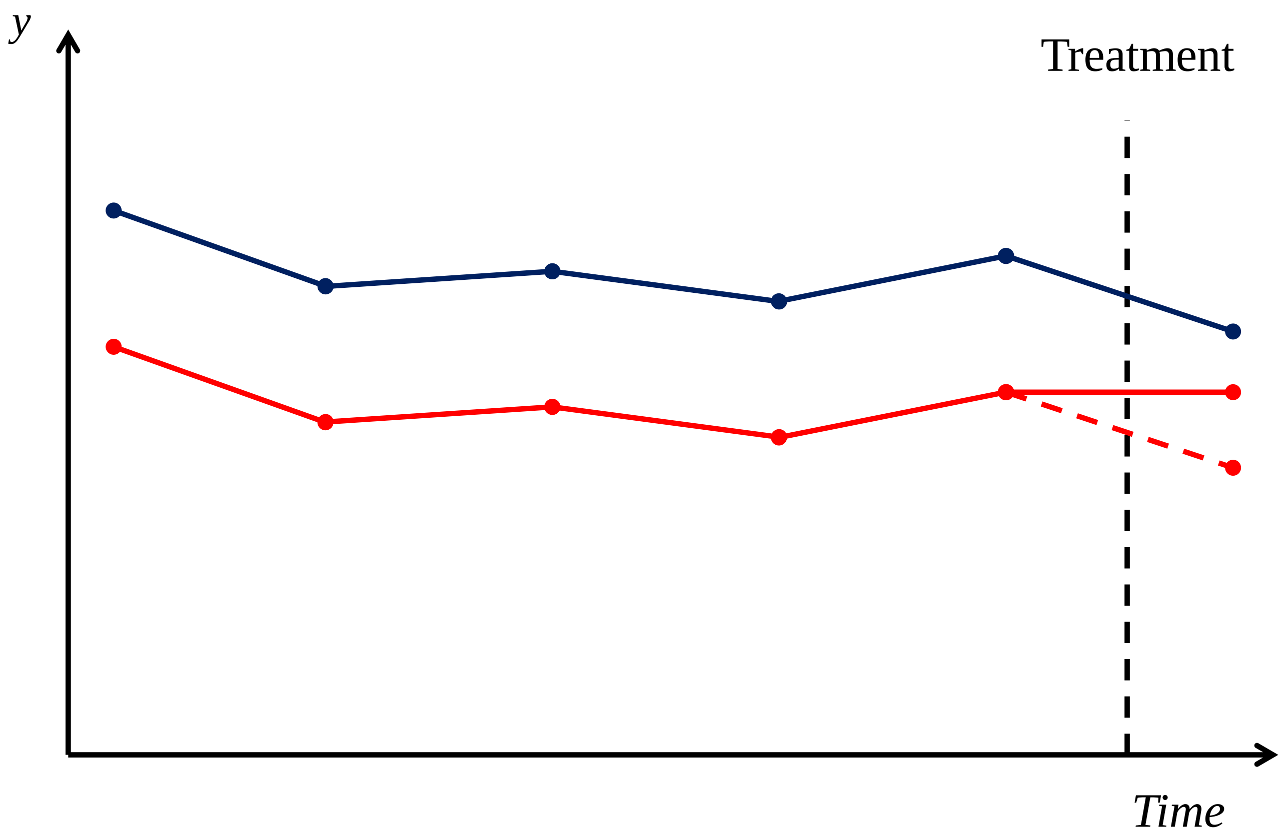
- **Surprisingly, employment rose in New Jersey!**

General case

Key assumption I

- **The key assumption for any DID strategy is that the outcome in treatment and control groups would follow the same time trend in the absence of the treatment**
  - This does not mean that they must have the same mean (or level) of the outcome!
- **This common or parallel trend assumption is impossible to test because you never observe the counterfactual**
  - But you can test it indirectly using pre-treatment data to show that the trends have been the same in past (only indirect evidence)
  - Even if pre-trends are the same one still must worry about other policies or something else changing at the same time

Common pre-treatment trends



Key assumption II

- **The second key assumption is that there are no spillover effects or that group compositions do not change because of treatment (if using repeated cross-sections)**
  - In the minimum wage example, this would mean that New Jersey's minimum wage increase does not directly affect employment in Pennsylvania
  - How plausible is this?

Key assumption III

- **The third key assumption is that nothing else happens at the same time as the treatment takes place that would affect the control and treatment groups differently**

Applications

ELSEVIER

European Journal of Political Economy

Volume 38, June 2015, Pages 140-152

Common pool problems in voluntary municipal mergers

Tuukka Saarimaa, Janne Tukiainen[1]

Show more

https://doi.org/10.1016/j.ejpoleco.2015.02.006

Get rights and content

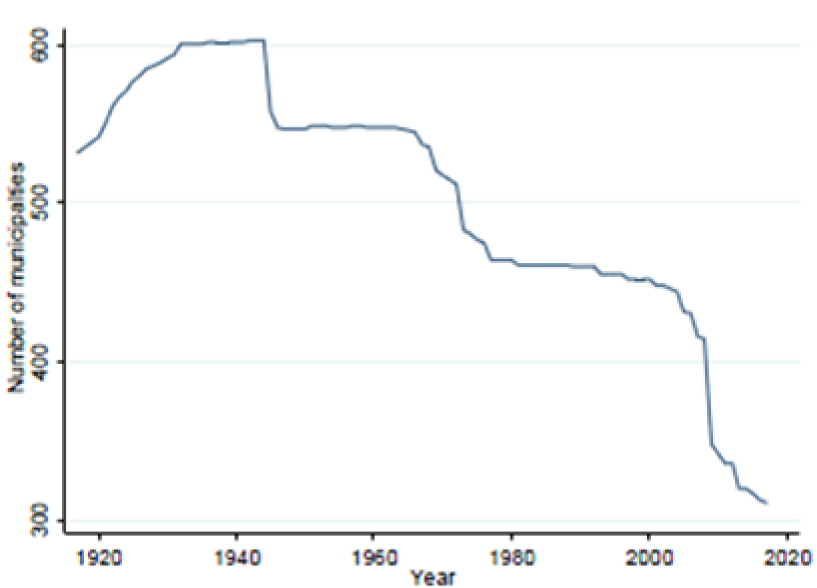
https://www.sciencedirect.com/science/article/abs/pii/S017626801500021X

Saarimaa & Tukiainen (2015)

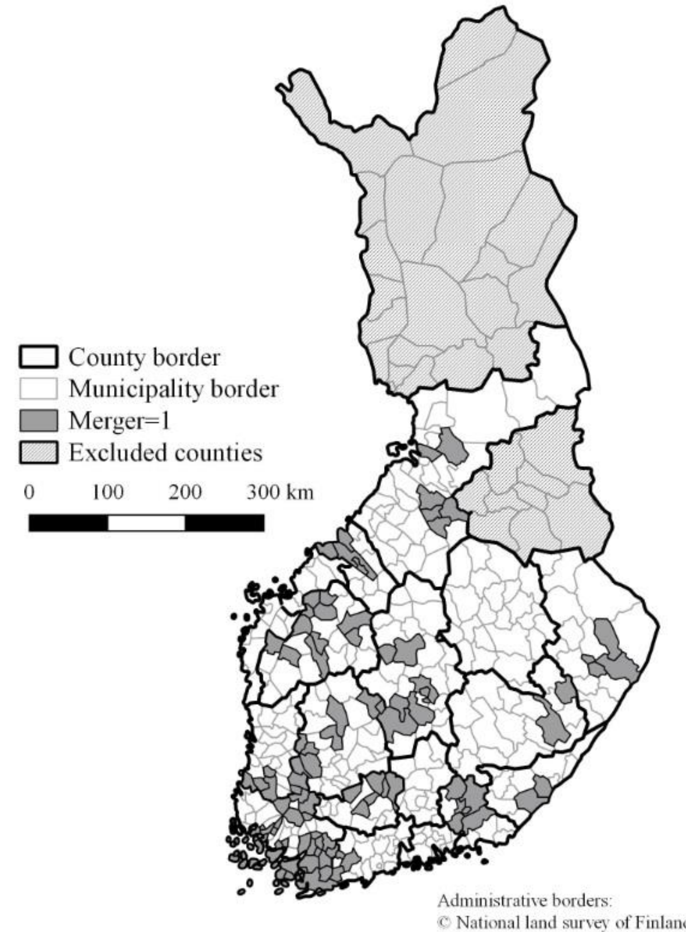
Highlights

- We analyze free-riding behavior of Finnish municipalities prior to municipal mergers.
- A time lag between the initial decision and the actual merger creates a common pool.
- Municipalities exploit the common pool by substantially increasing municipal debt.
- The results are consistent with the “law of $1/n$”.

The mergers



600
500
400
300
Number of municipalities
1920
1940
1960
1980
2000
2020
Year



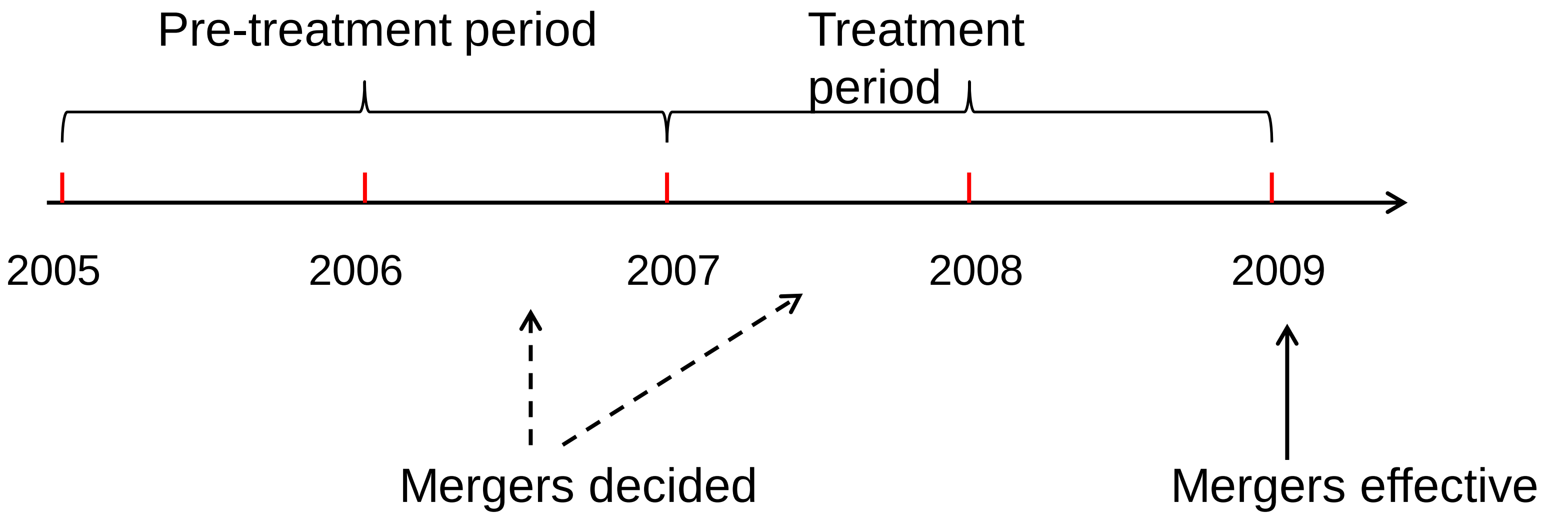
County border
Municipality border
Merger=1
Excluded counties
0
100
200
300 km
Administrative borders:
© National land survey of Finland

Motivation

- **A common pool problem arises when the costs of an activity that benefits a small group of people are shared among a larger group**
  - "Law of 1/*n*" (Weingast et al. 1981, JPE): Due to common pool funding and universalism, spending increases as the number of decision makers increase => free-riding
- **Municipality mergers open an opportunity to study the common pool problem and the law of 1/n**
  - Municipalities remain autonomous before the merger
  - Free-riding incentives related to relative size of merging municipalities, not the number of municipalities in the merger
  - Electoral punishment unlikely

The mergers

- **In 2007 a provisional law introduced merger subsidies and other merger incentives**
- **At the start of 2009, 32 mergers (involving 99 municipalities) took place; decided in 2006–07**



Common pool incentives

- **Who has incentives to free-ride?**
- **Some incentives for all municipalities that merge, but stronger for relatively small municipalities**
  - We define a measure of free-riding incentives for municipality i in merger j as

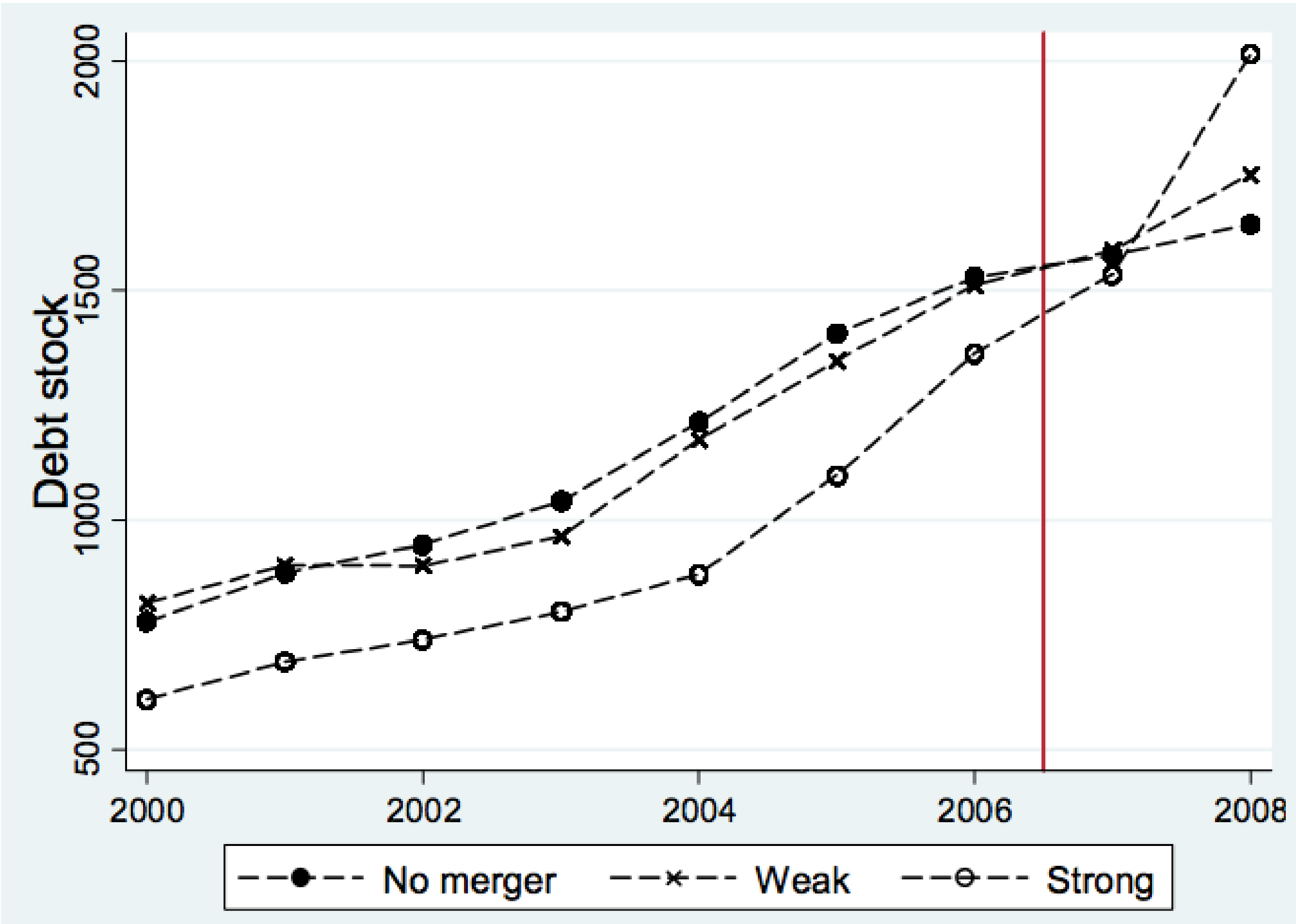
    $$freeride_i = 1 - taxbase_i / taxbase_j$$

  - Idea: municipality i internalizes $taxbase_i / taxbase_j$ of the social marginal cost of borrowing
  - Higher values of *freeride* imply stronger incentives to freeride

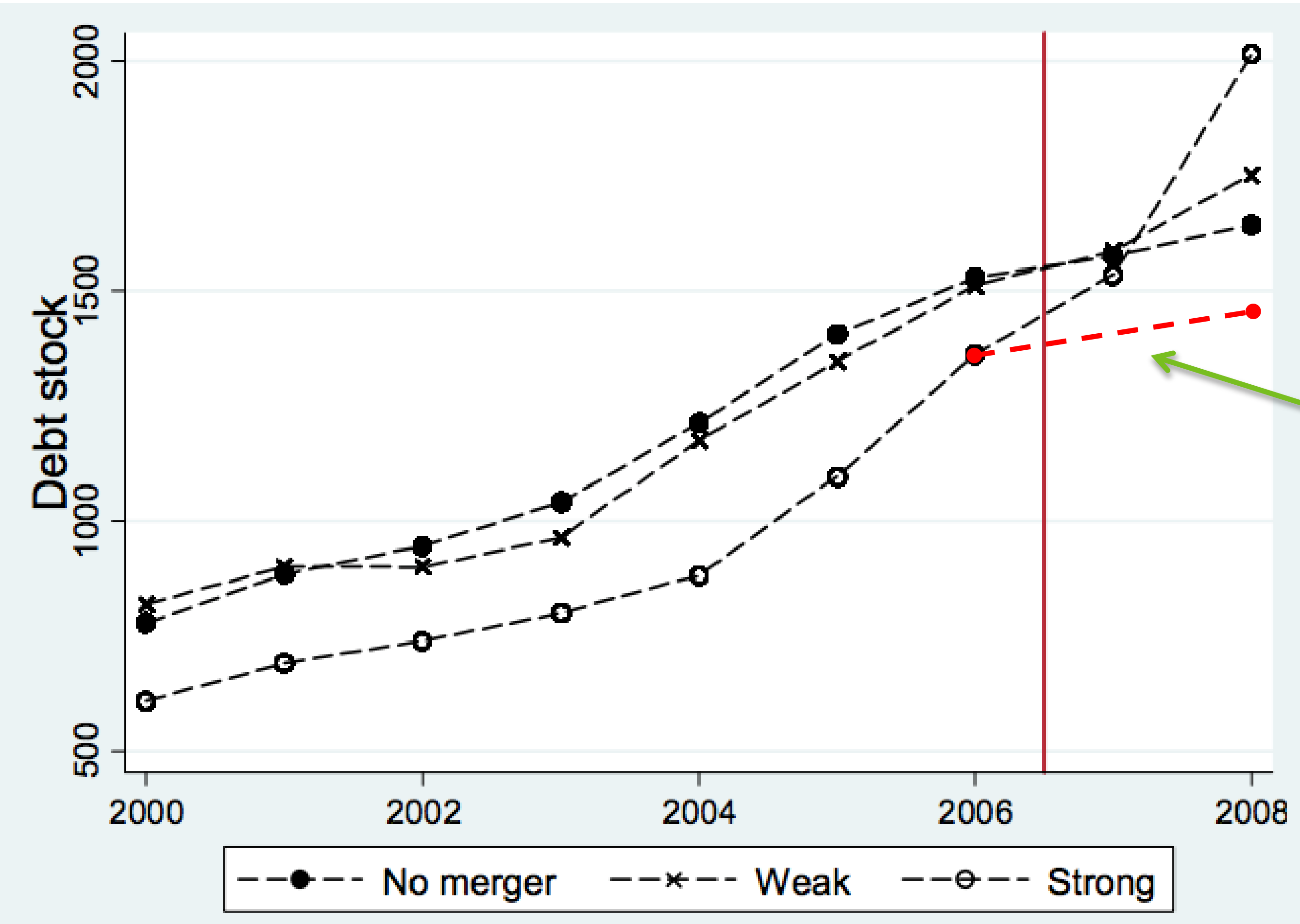
Analysis

- **The incentives measure is continuous, so for the graphical analysis we use groups**
  - Divide merger group into municipalities with weak and strong incentives to free-ride (according to median value of *freeride*)
  - Compare these groups to the no-merger group and each other
  - So, we have two treatment groups and a control group, but the DID idea is the same

Main results

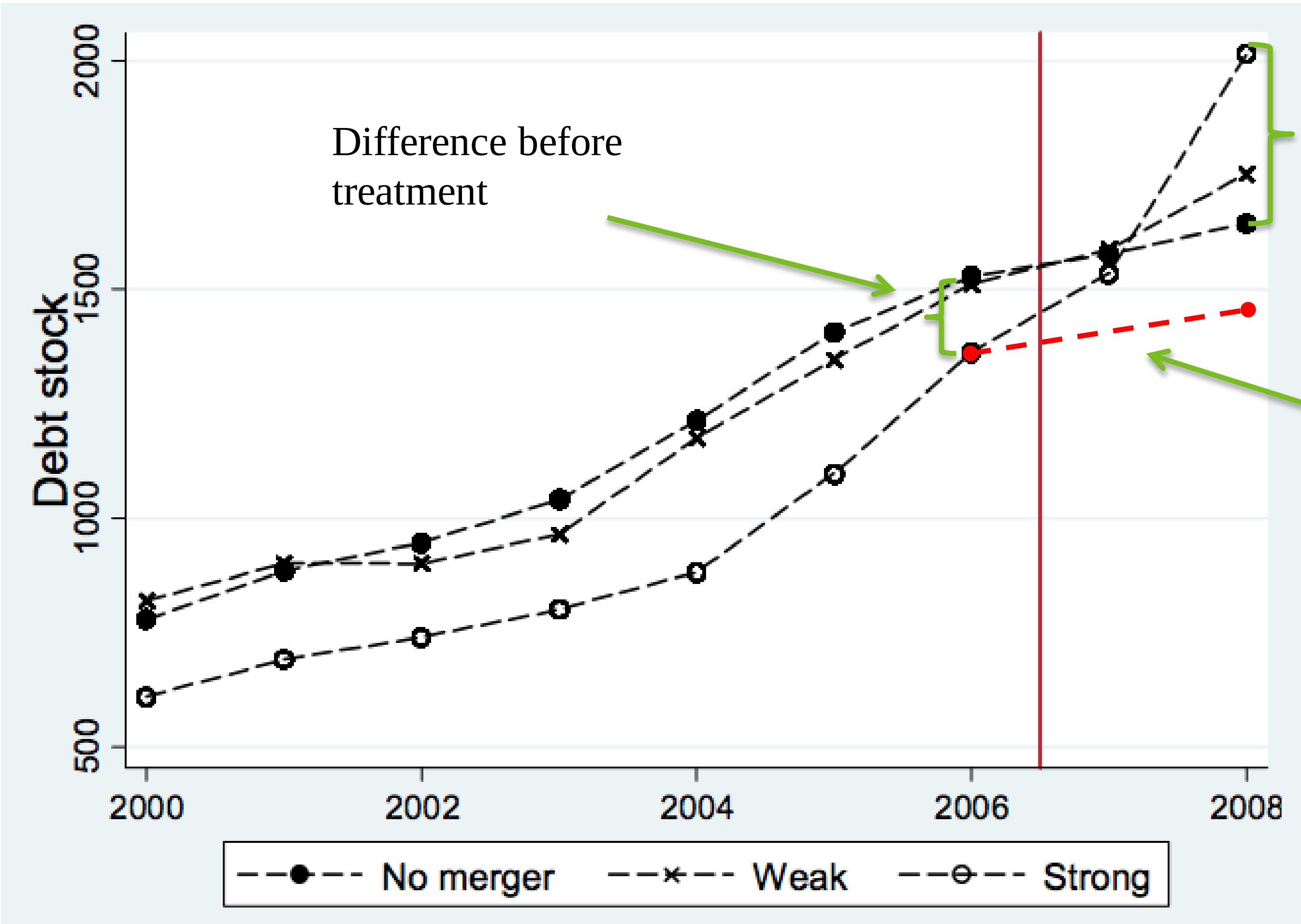


Main results



Counterfactual trend in the treatment groups

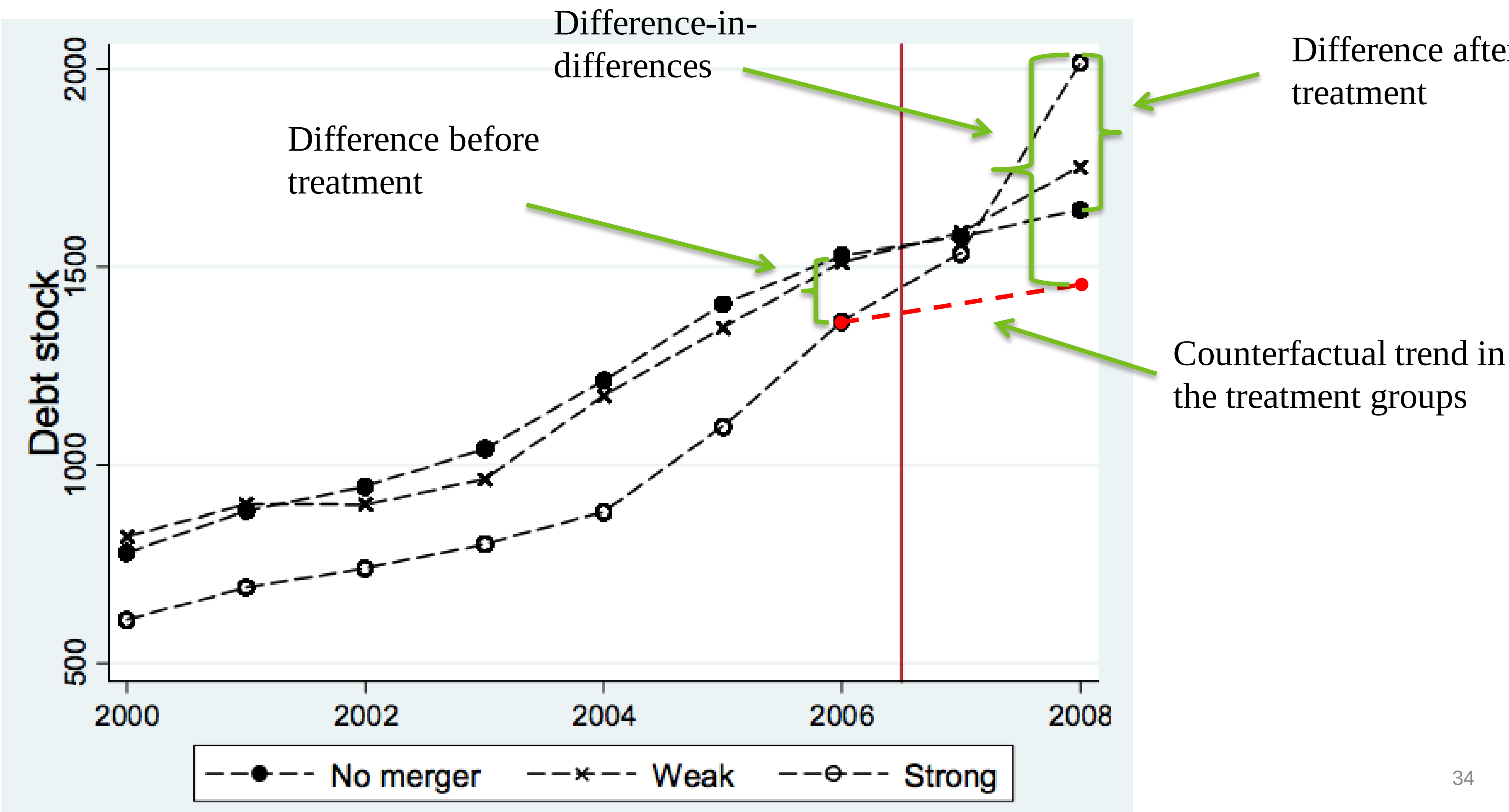
Main results



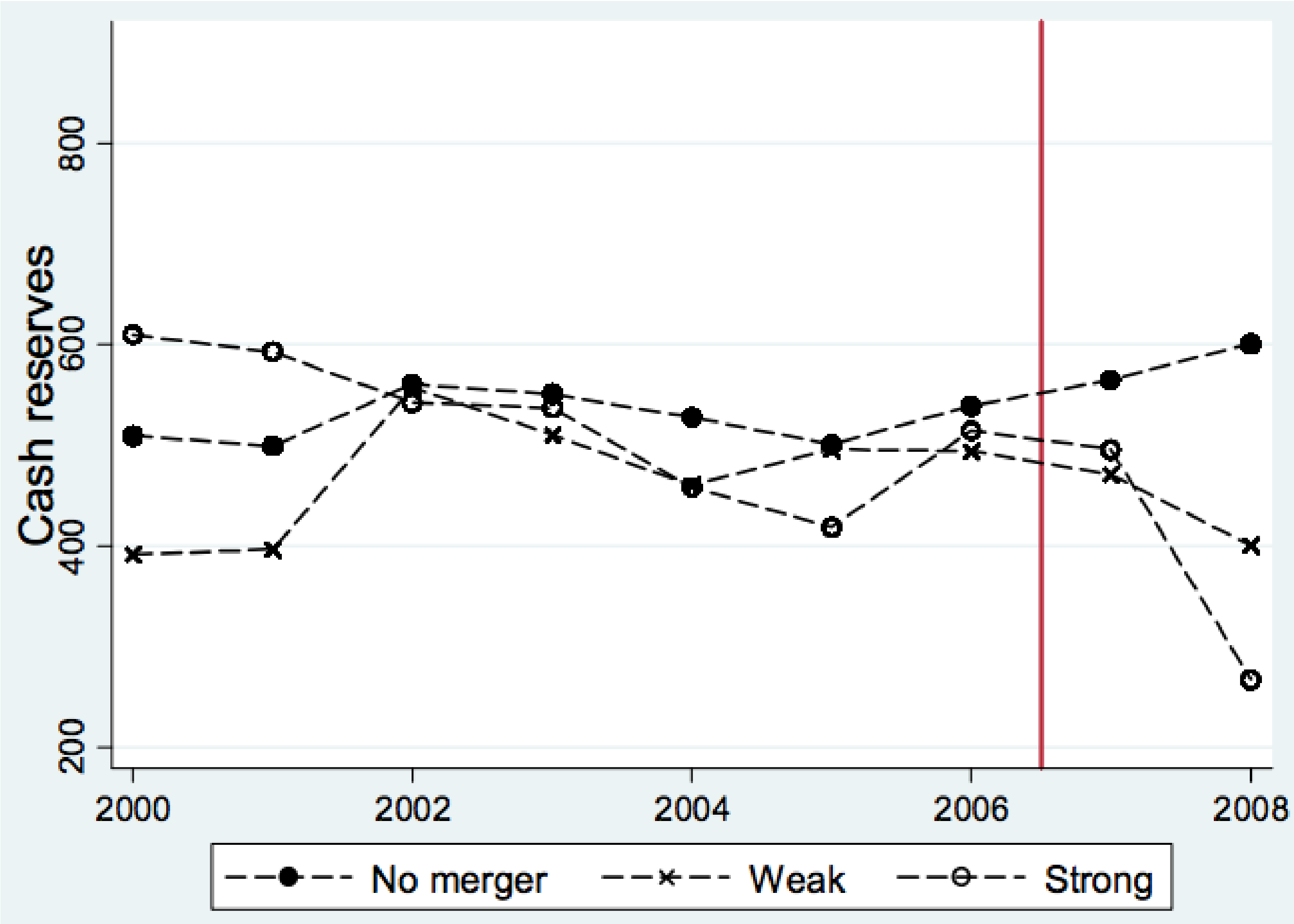
Difference after
treatment

Counterfactual trend in
the treatment groups

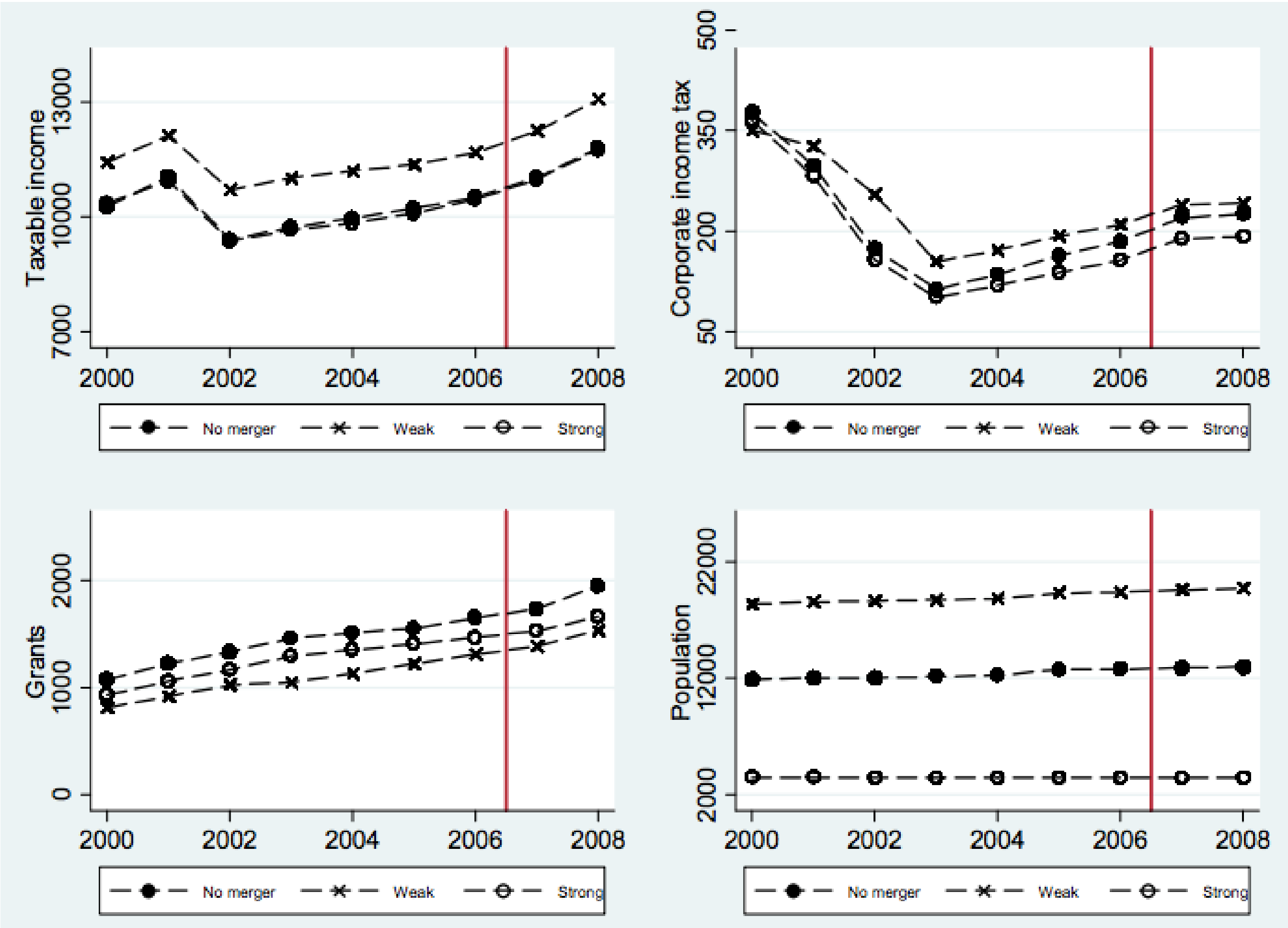
Main results



Main results

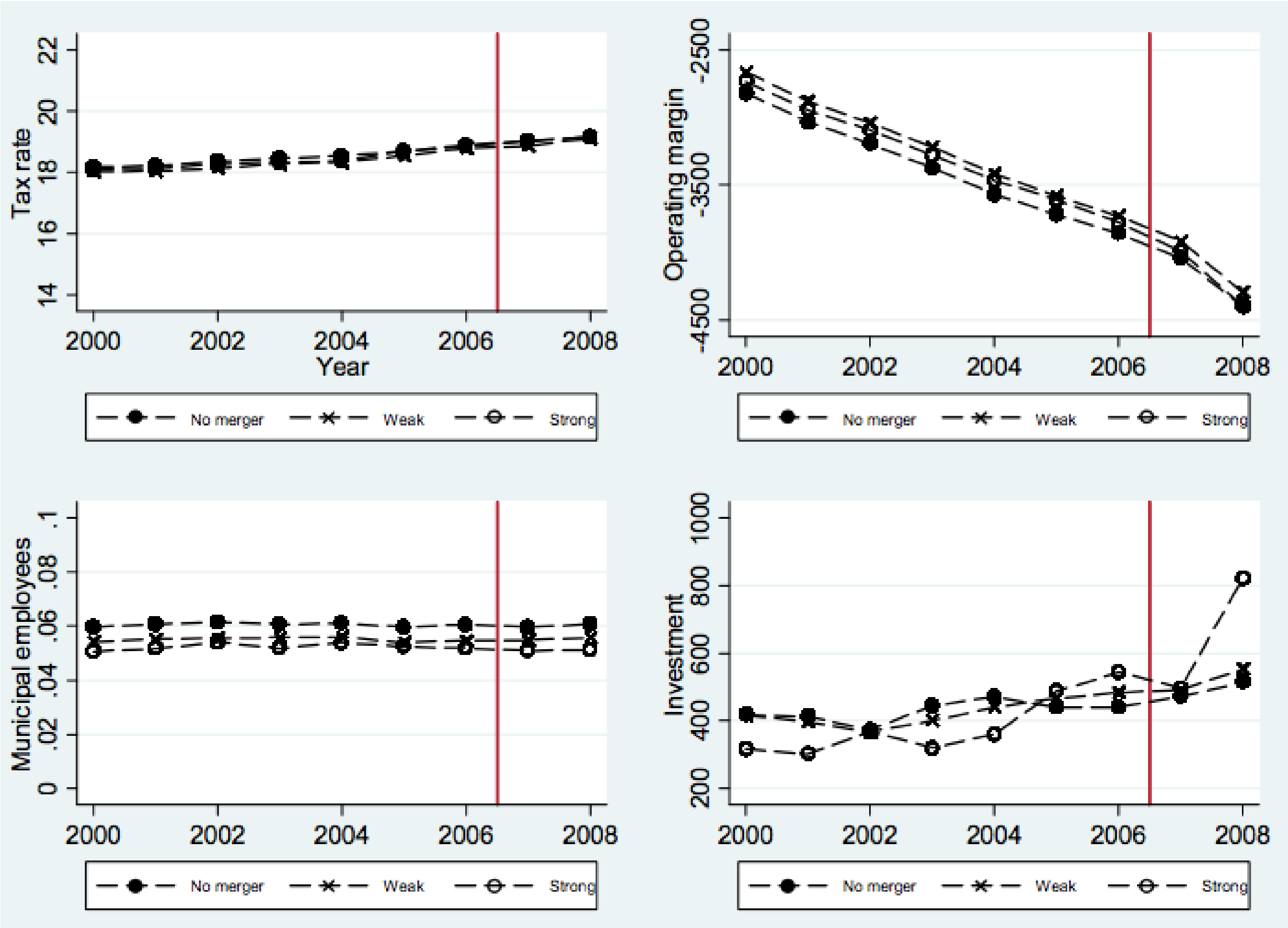


Other explanations

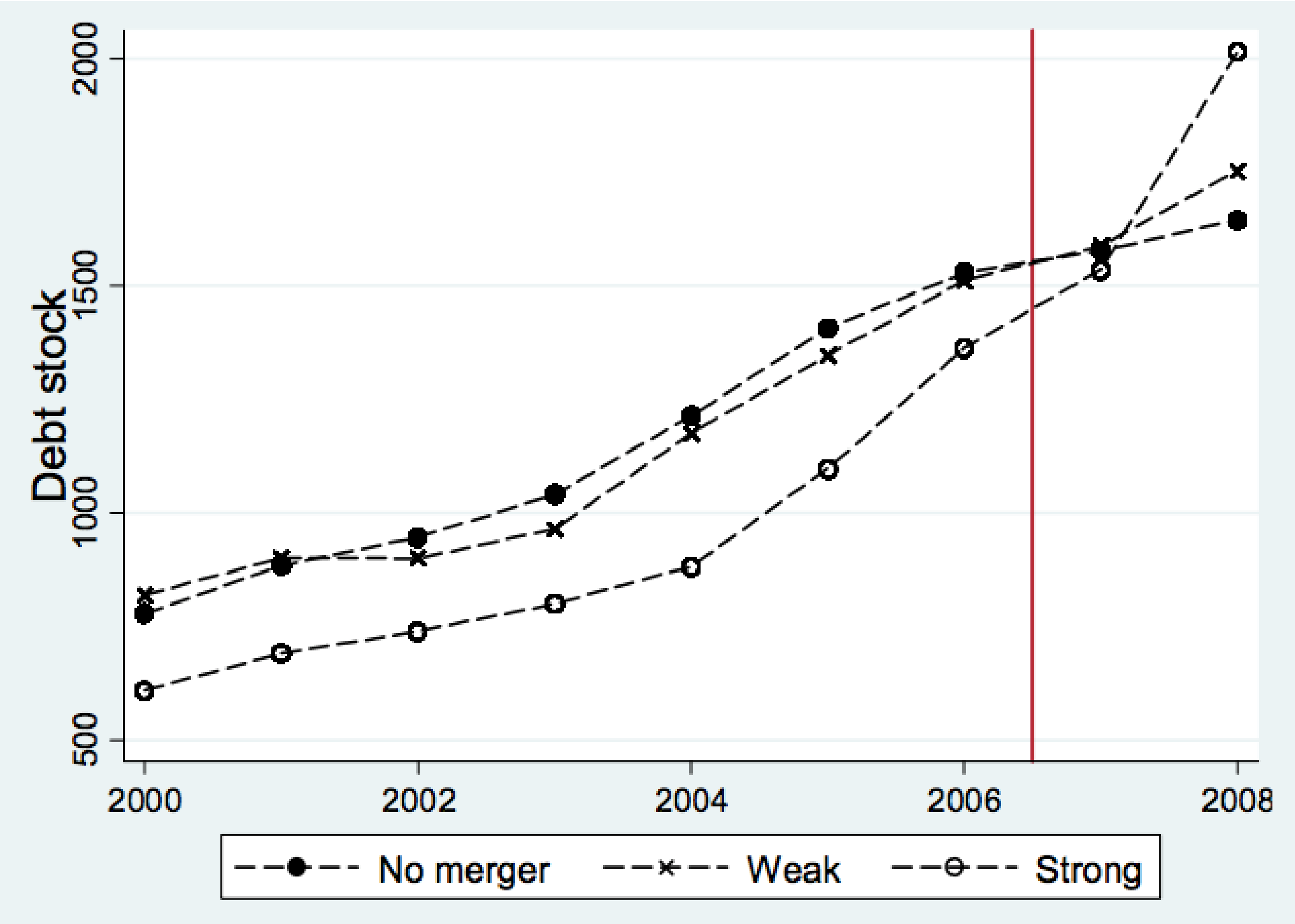


Did something else change at the same time that could explain the increase in debt?

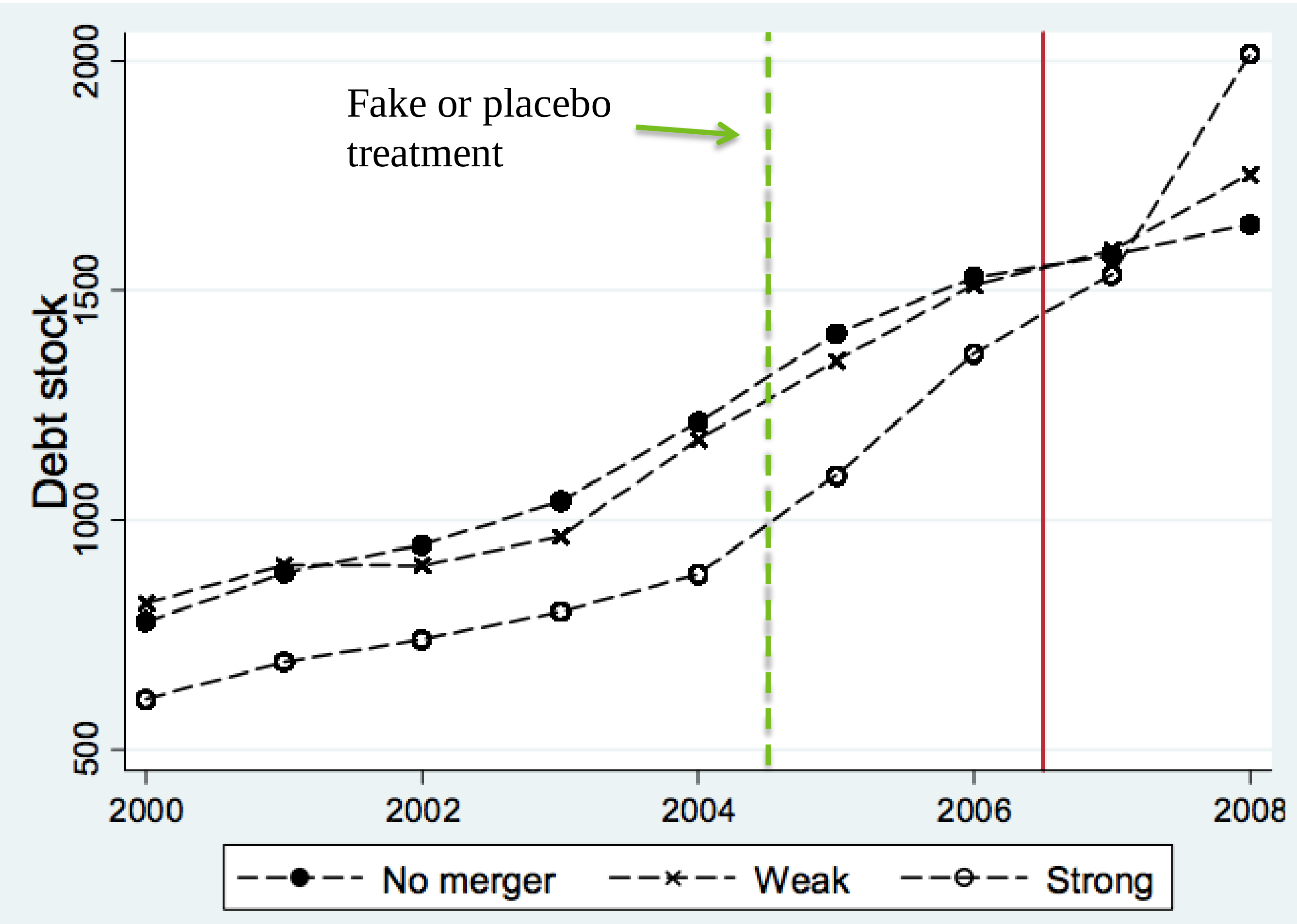
Show me the money



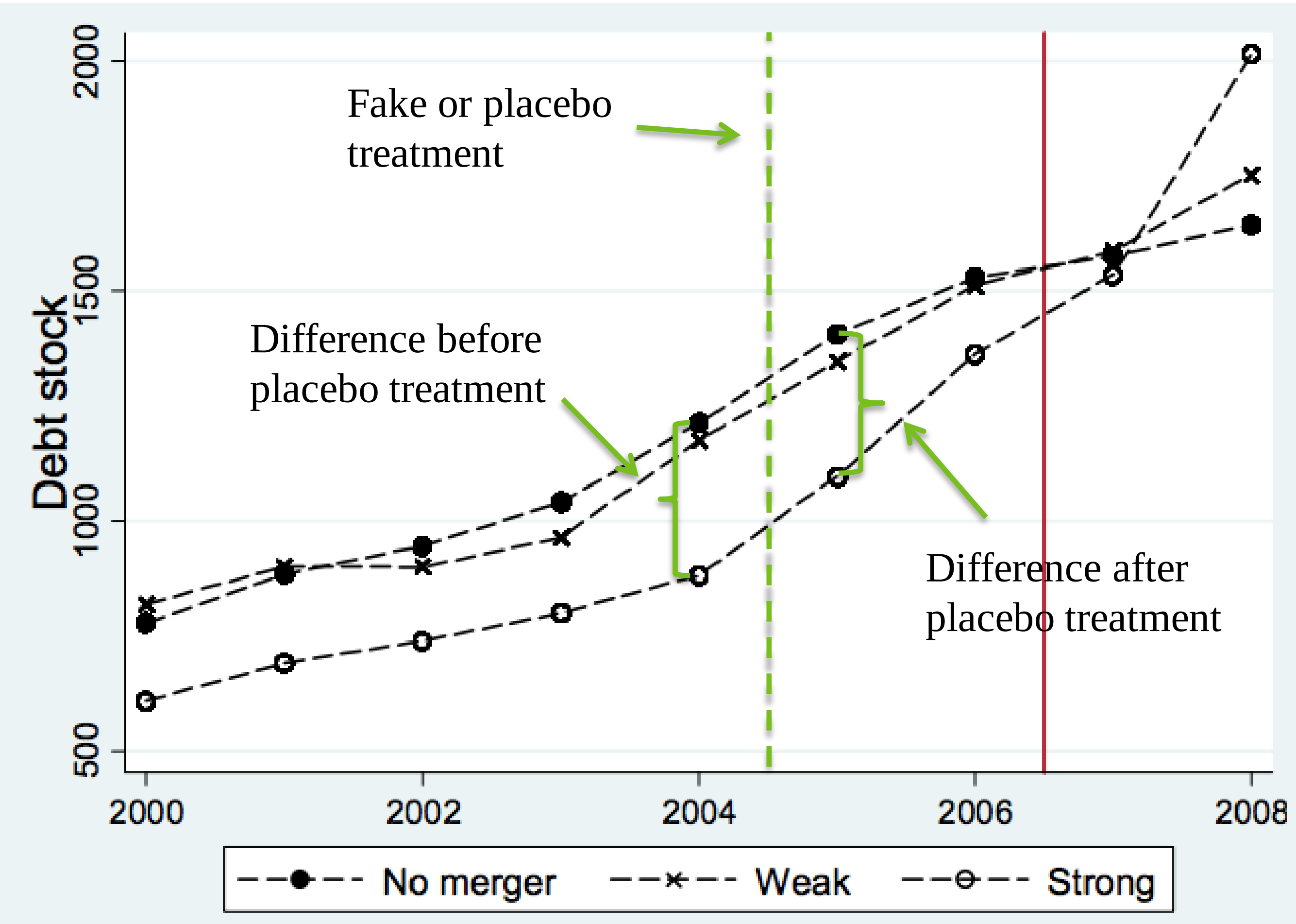
Testing for common trends (indirectly)



Testing for common trends (indirectly)



Testing for common trends (indirectly)



Estimate a DID as if the treatment took place at the placebo cutoff!

Repeat with all the pre-treatment years!

If you have a lot of pre-treatment data, showing placebo tests is a convincing way to argue that the common trends assumption is valid

Conclusions – Saarimaa & Tukiainen (2015)

- **Consistent with the law of 1/*n*, find clear evidence of free-riding among merging Finnish municipalities**
- **Some policy advice:**
  - Policy 1: Politicians are likely to exploit a common pool if given the opportunity to do so (similar evidence from other countries)
  - Policy 2: During a merger process, some financial constraints on the local level may be a good idea (as in Denmark)
  - Policy 3: For the mergers to be beneficial overall, benefits need to be fairly large

American Economic Journal: Applied Economics 3 (January 2011): 65–90
http://www.aeaweb.org/articles.php?doi=10.1257/app.3.1.65

Traffic Congestion and Infant Health: Evidence from E-ZPass†

By Janet Currie and Reed Walker*

We exploit the introduction of electronic toll collection, (E-ZPass), which greatly reduced both traffic congestion and vehicle emissions near highway toll plazas. We show that the introduction of E-ZPass reduced prematurity and low birth weight among mothers within 2 kilometers (km) of a toll plaza by 10.8 percent and 11.8 percent, respectively, relative to mothers 2–10 km from a toll plaza. There were no immediate changes in the characteristics of mothers or in housing prices near toll plazas that could explain these changes. The results are robust to many changes in specification and suggest that traffic congestion contributes significantly to poor health among infants. (*JEL* I12, J13, Q51, Q53, R41)

https://www.aeaweb.org/articles?id=10.1257/app.3.1.65

Motivation

- **Motor vehicles are a major source of air pollution**
  - Nationally they are responsible for over 50 percent of carbon monoxide (CO), 34 percent of nitrogen dioxide (NO2), and over 29 percent of hydrocarbon emissions, in addition to as much as 10 percent of fine particulate matter emissions
- **In urban areas, vehicles are the dominant source of these emissions**

Research question

- **Studies the effect of E-ZPass, and thus the sharp reductions in local traffic congestion, on the health of infants born to mothers living near toll plazas**
- **This is interesting for three reasons:**
  1. There is increasing evidence of the long-term effects of poor health at birth on future outcomes
  2. The study of newborns overcomes several difficulties in making the connection between pollution and health because the link between cause and effect is immediate.
  3. E-ZPass is an interesting policy experiment because, while pollution control was an important consideration for policy makers, the main motive for consumers to sign up for E-ZPass is to reduce travel time

Selection bias

- **Since air pollution is not randomly assigned, studies comparing health outcomes for populations exposed to differing pollution levels may not adequately control for confounding determinants of health**
  - Families with higher incomes or preferences for cleaner air are likely to sort into locations with better air quality, and failure to account for this sorting will lead to overestimates of the effects of pollution
  - Alternatively, pollution levels are higher in urban areas where there are often more educated individuals with better access to health care, which can cause underestimates of the true effects of pollution on health

Empirical strategy

- **In the absence of a randomized trial, we**
  - exploit a policy change that created large local and persistent reductions in traffic congestion and traffic related air emissions for certain segments along a highway
  - compare the infant health outcomes of those living near an electronic toll plaza before and after implementation of E-ZPass to those living near a major highway but further away from a toll plaza

Empirical strategy

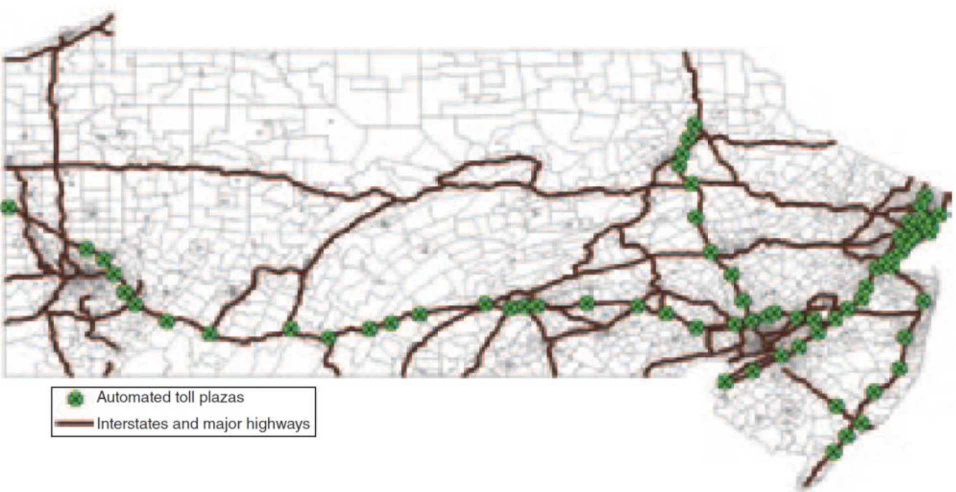


Figure 1. Locations of Toll Plazas and Major Roadways in New Jersey and Pennsylvania

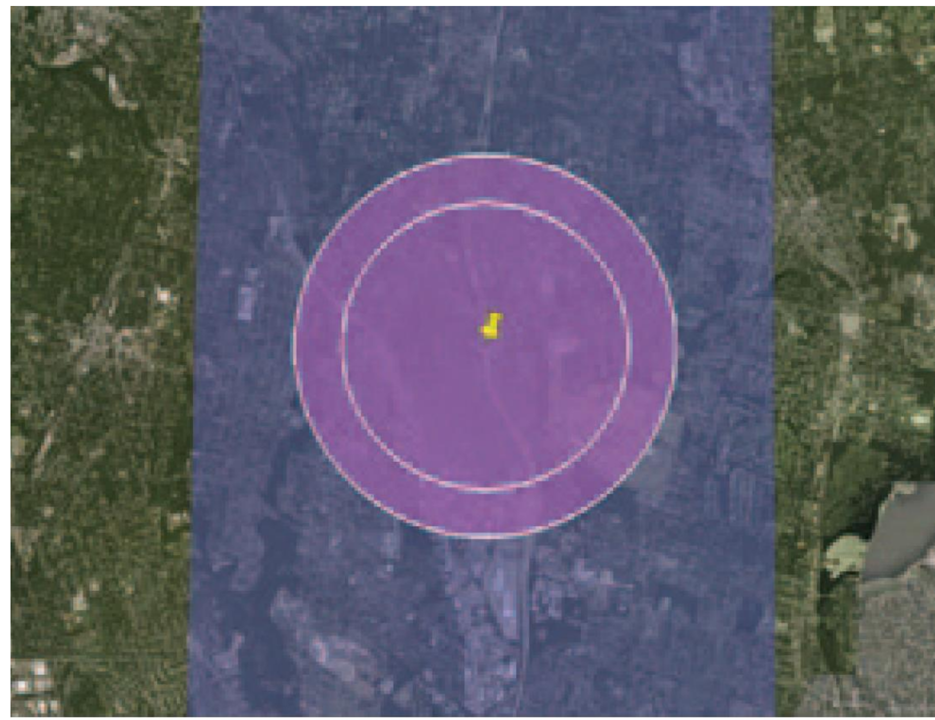


Figure 2. Research Design Showing 1.5 km and 2 km Treatment Radii and 3 km from Highway Control Group

Empirical strategy

- **“Specifically, we compare mothers within 2 km of a toll plaza to mothers who are between 2 km and 10 km from a toll plaza, but still within 3 km, of a major highway before and after the adoption of E-ZPass in New Jersey and Pennsylvania.”**
- **Assumption:**
  - “Our difference in differences research design relies on the assumption that the characteristics of mothers near a toll plaza change over time in a way that is comparable to those of other mothers who live further away from a plaza, but still close to a major highway.”

Results – desc stat and crude DID

TABLE 1—SUMMARY STATISTICS

	<2 km E-ZPass before	<2 km E-ZPass after	>2 km and <10 km E-ZPass before	>2 km and <10 km E-ZPass after	>10 km Toll plaza
Panel A. Difference-in-difference sample					
Outcomes					
Premature	0.095	0.095	0.102	0.109	0.085
Low birth weight	0.082	0.078	0.089	0.092	0.078
Controls					
Mother Hispanic	0.291	0.332	0.165	0.229	0.054
Mother black	0.16	0.173	0.233	0.264	0.047
Mother education	13.12	13.2	13.276	13.24	12.92
Mother HS dropout	0.169	0.164	0.154	0.163	0.173
Mother smoked	0.089	0.075	0.109	0.086	0.152
Teen mother	0.073	0.061	0.082	0.069	0.079
Birth order	1.3	1.37	1.39	1.46	1.68
Multiple birth	0.028	0.033	0.032	0.037	0.033
Child male	0.51	0.512	0.514	0.512	0.512
Distance to roadway	1.099	1.074	1.507	1.482	21
Observations	33,758	29,677	190,904	161,145	185,795
NJ observations	26,415	26,563	128,547	133,560	70,484
PA observations	7,343	3,114	62,357	27,585	115,311

Main results

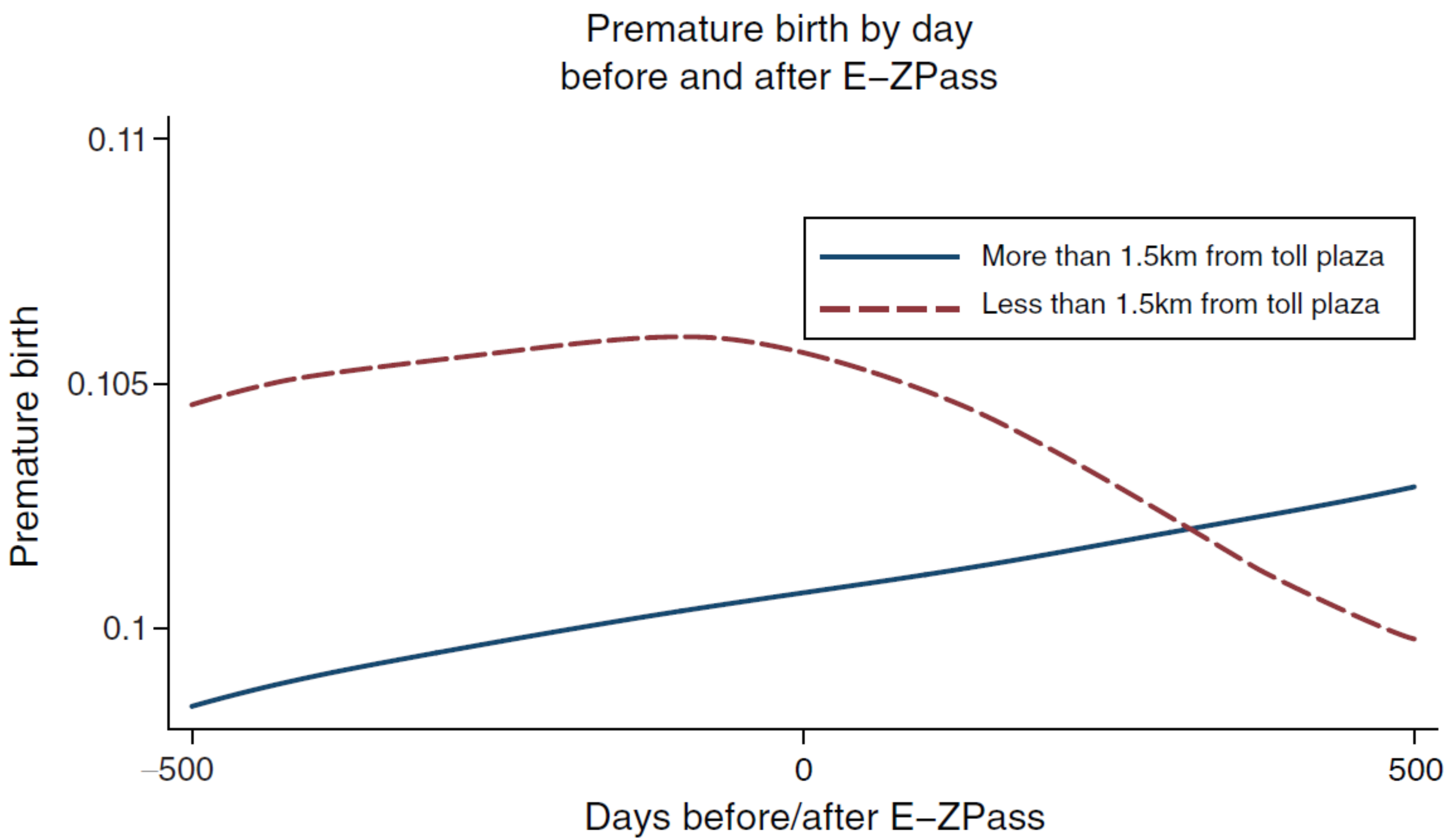


FIGURE 6

Main results

Counterfactual trend in the treatment groups

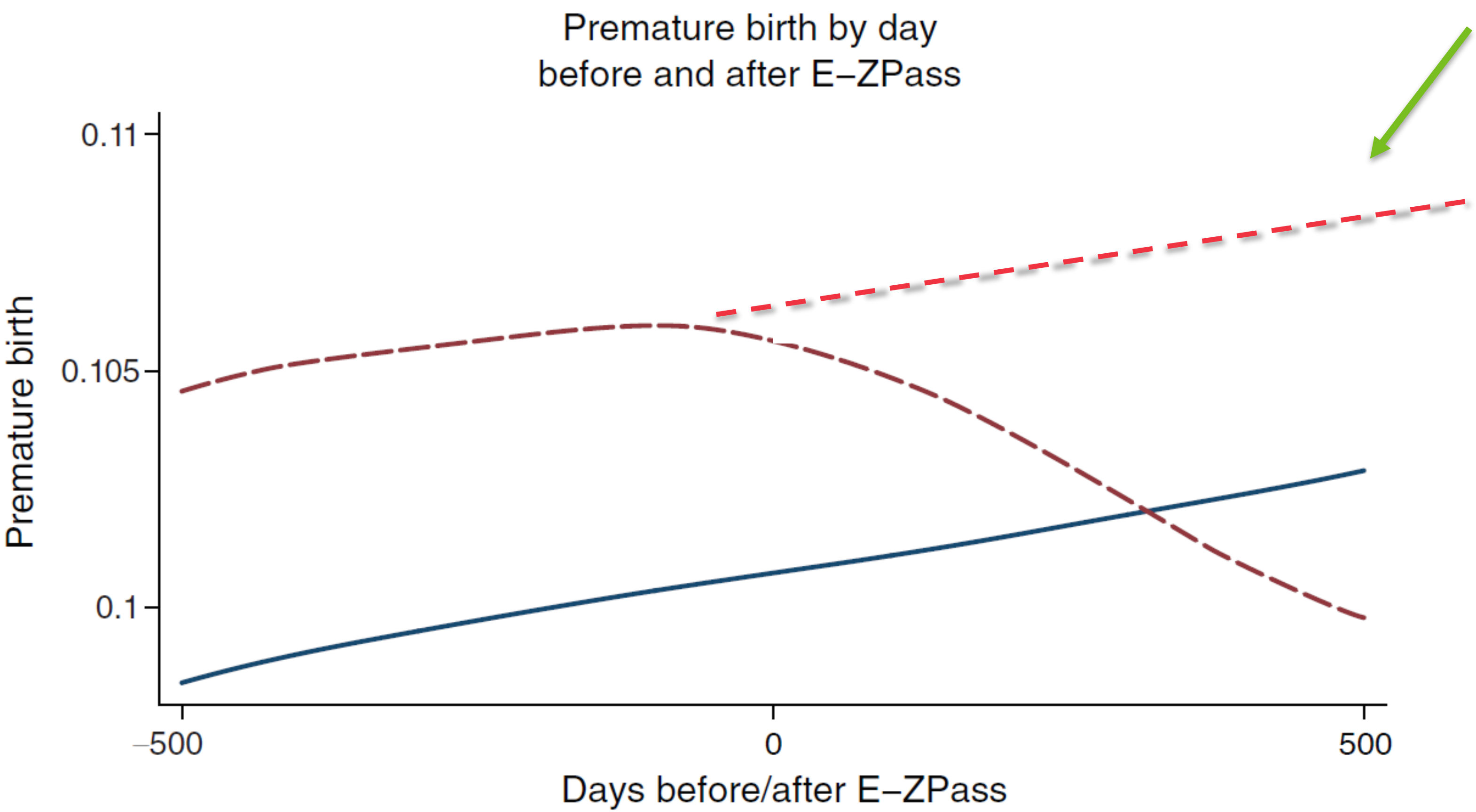


FIGURE 6

Main results

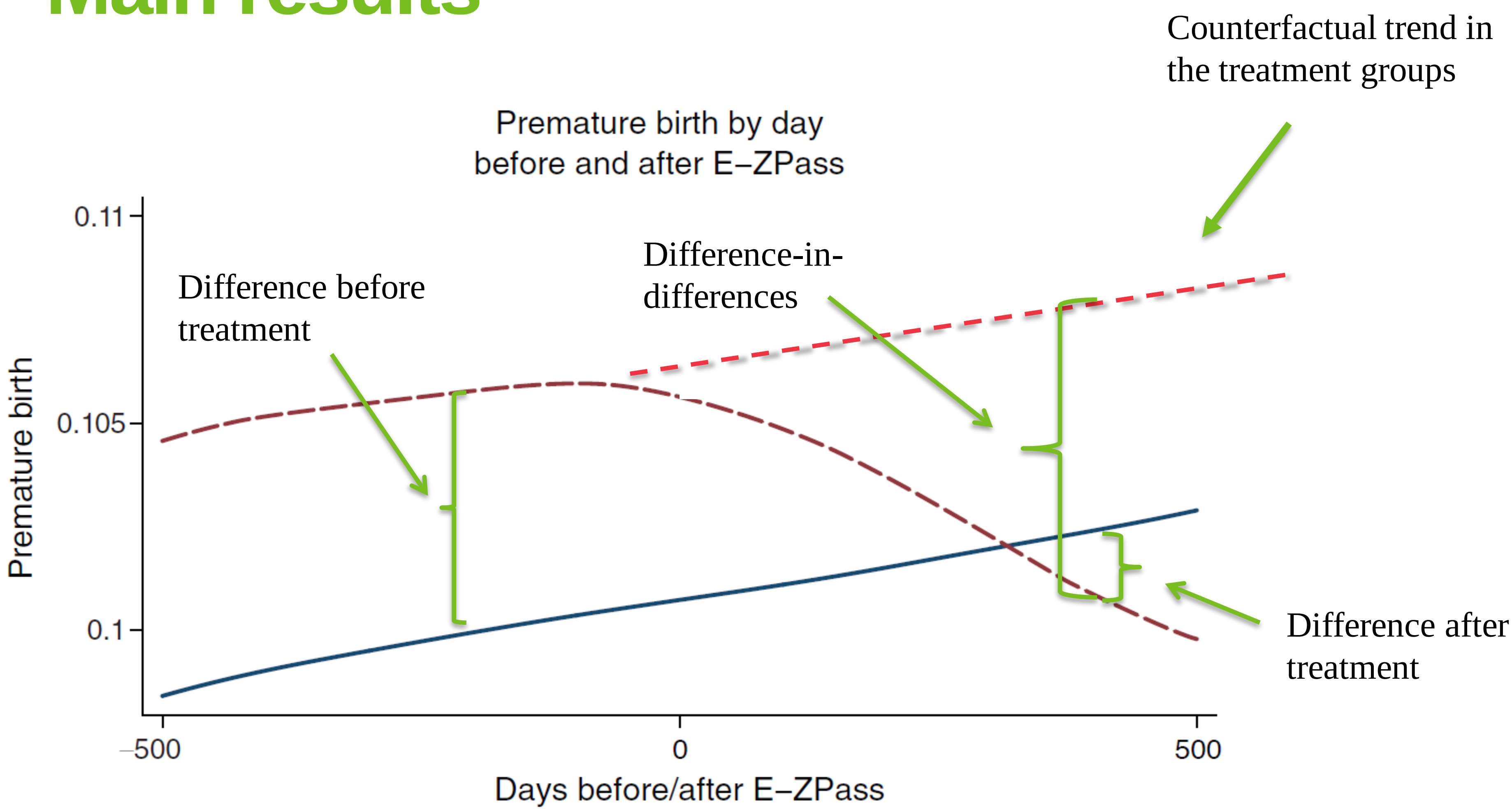


FIGURE 6

Main results

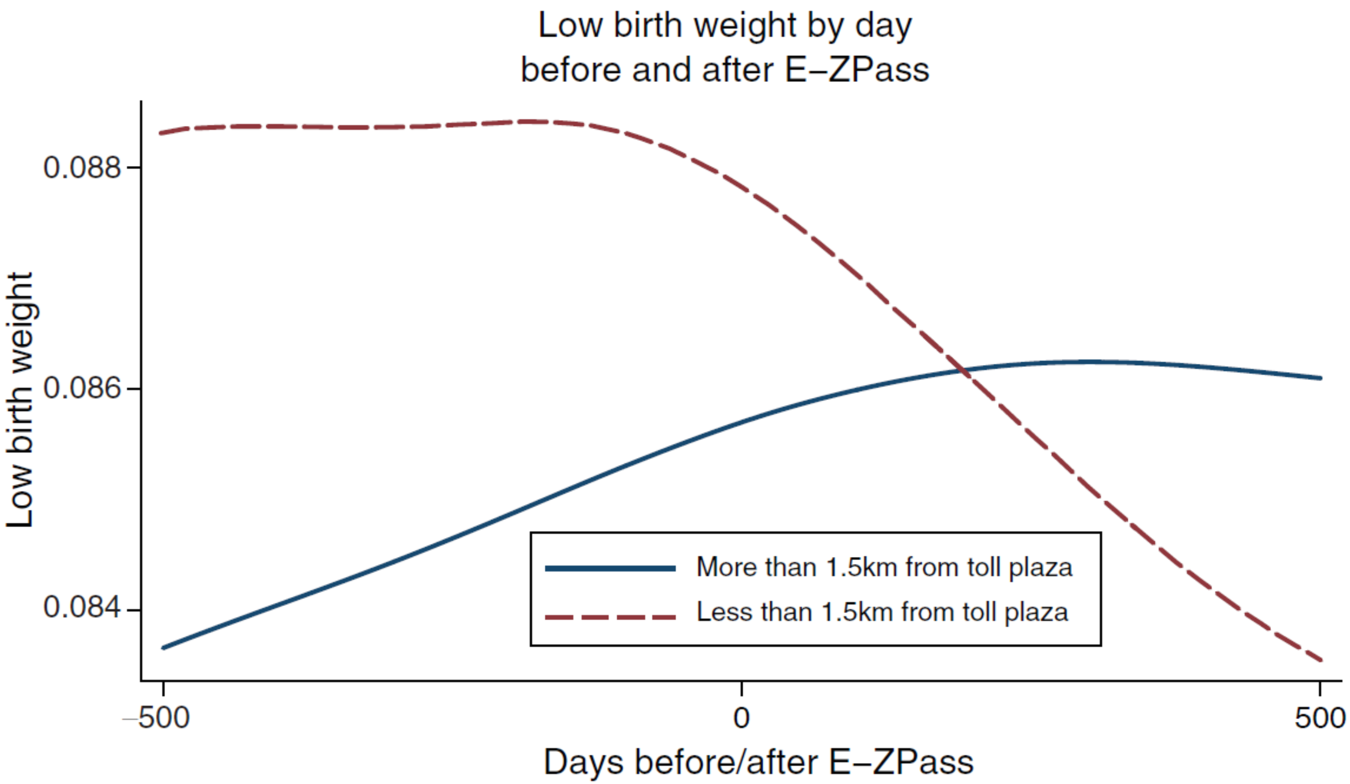


FIGURE 5

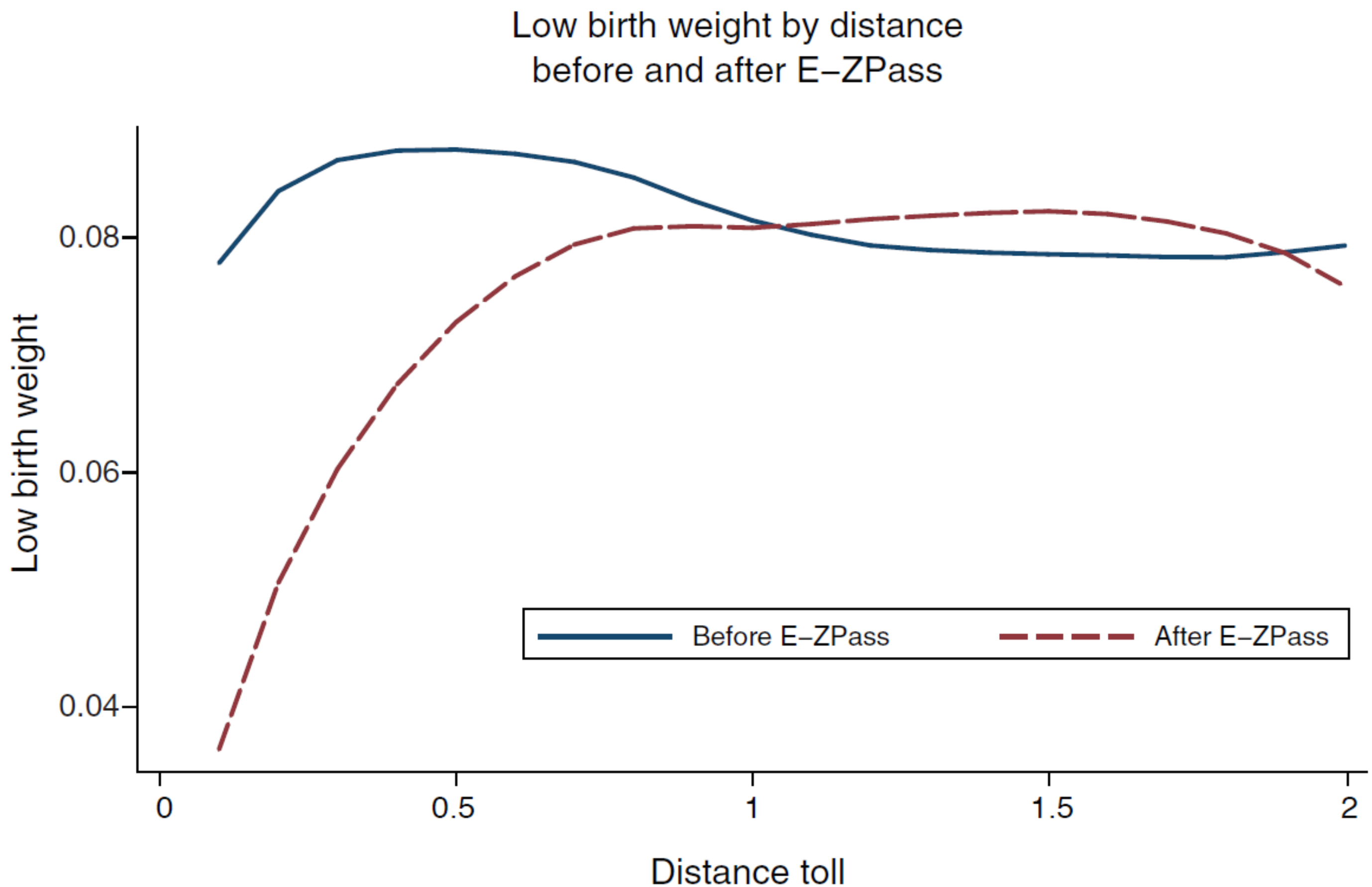


FIGURE 3

Notes: Smoothed plots of treatment and control groups using locally weighted regression. To facilitate computation, observations are first grouped into 0.1-mile bins by treatment and control and averaged. The weights are applied using a tricube weighting function (William S. Cleveland 1979) with a bandwidth of 1.

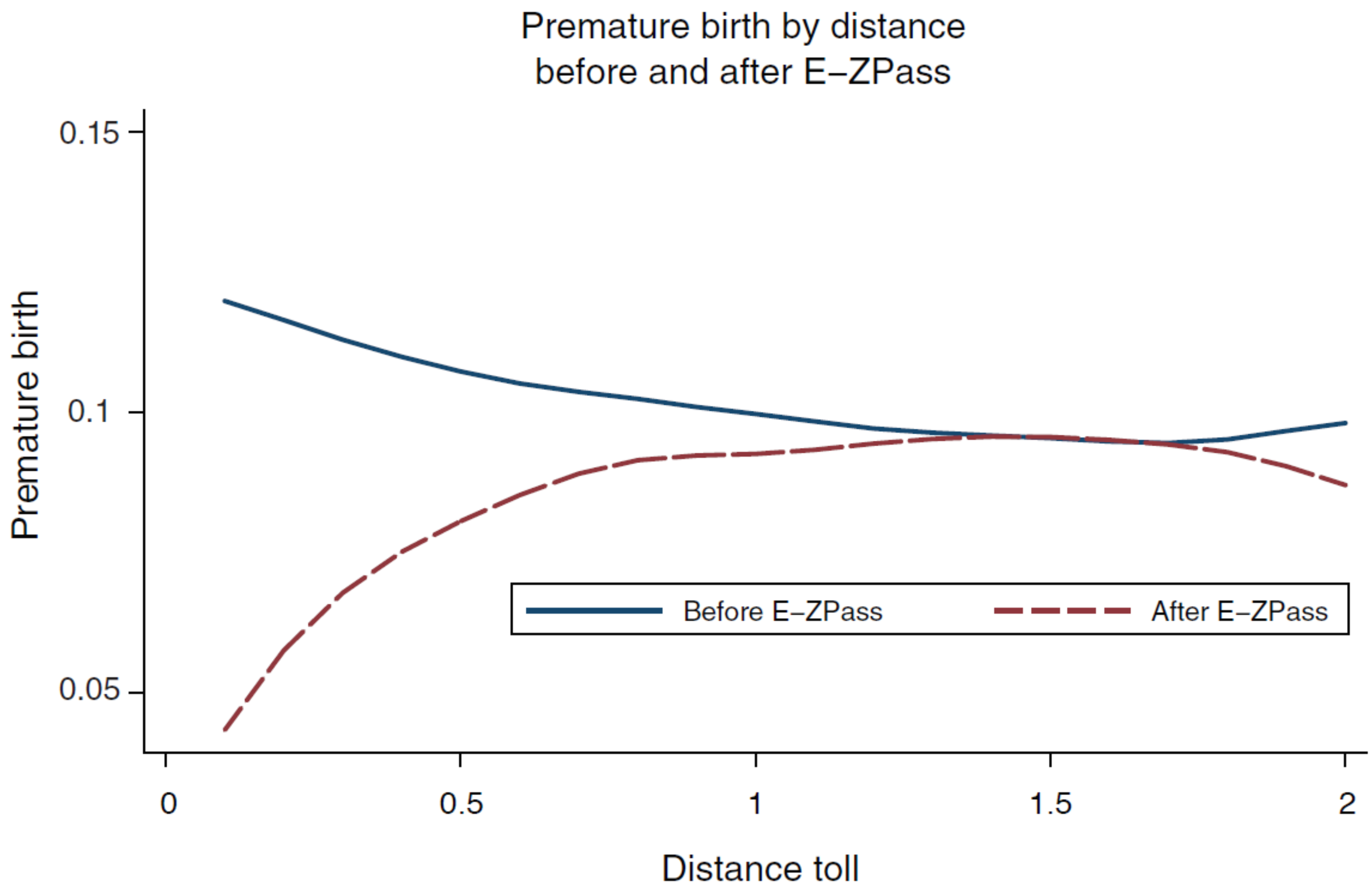


FIGURE 4

Notes: Smoothed plots of treatment and control groups using locally weighted regression. To facilitate computation, observations are first grouped into 0.1-mile bins by treatment and control and averaged. The weights are applied using a tricube weighting function (Cleveland 1979) with a bandwidth of 1.

Conclusions – Currie & Reed (2011)

- **We show that E-ZPass reduced the incidence of prematurity and low birth weight in the vicinity of toll plazas by 6.7–9.1 percent and 8.5–11.3 percent, respectively**
  - These are large but not implausible effects given the correlations between proximity to traffic and birth outcomes found in previous studies
- **Results suggest that policies intended to curb traffic congestion can have significant health benefits for local populations in addition to the more often cited benefits in terms of reducing travel costs**

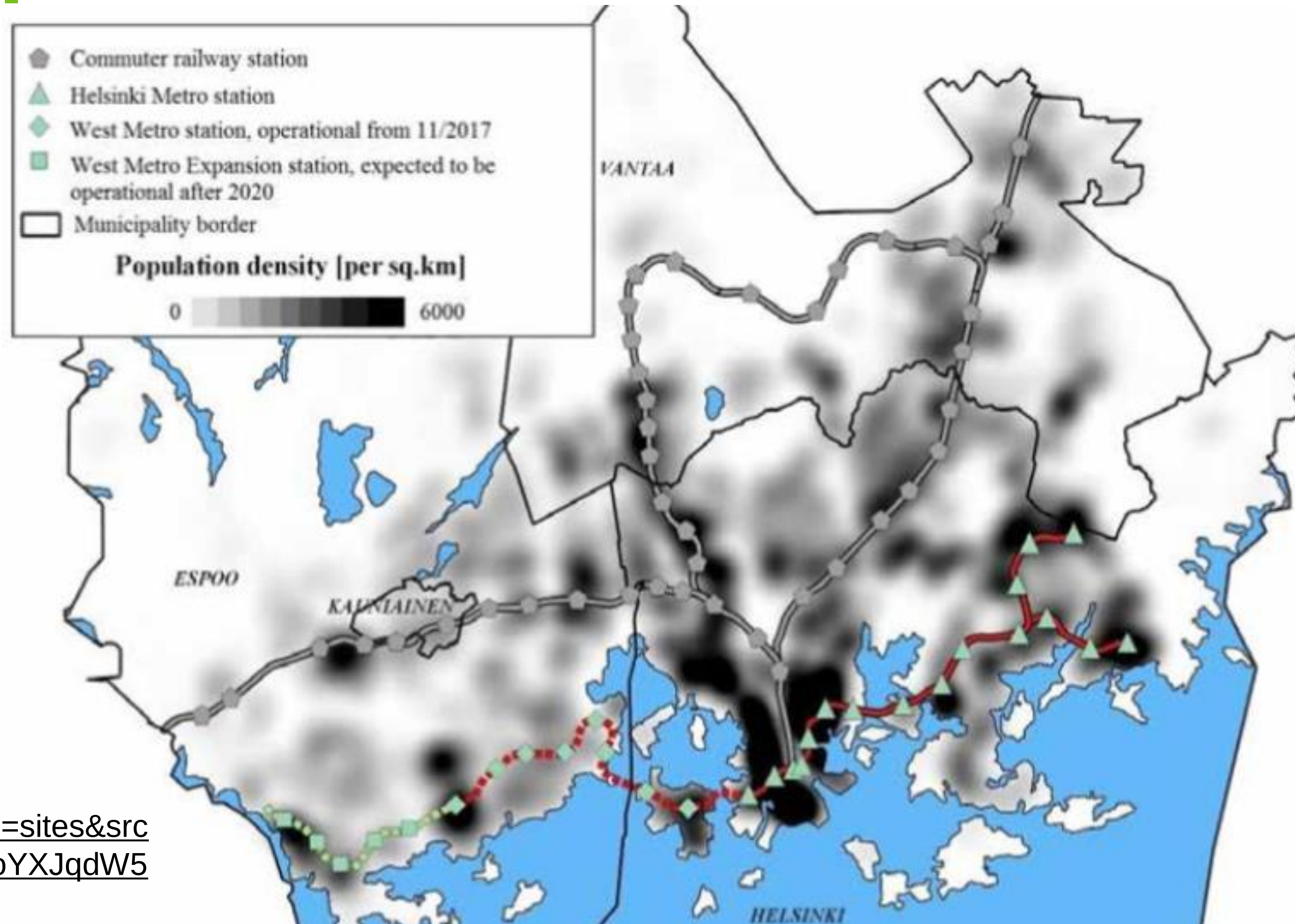
Other examples

Harjunen (2018): West Metro extension and house prices

Fig. A2. A route map of the metro in Helsinki in 2016 (© HSL 2016)

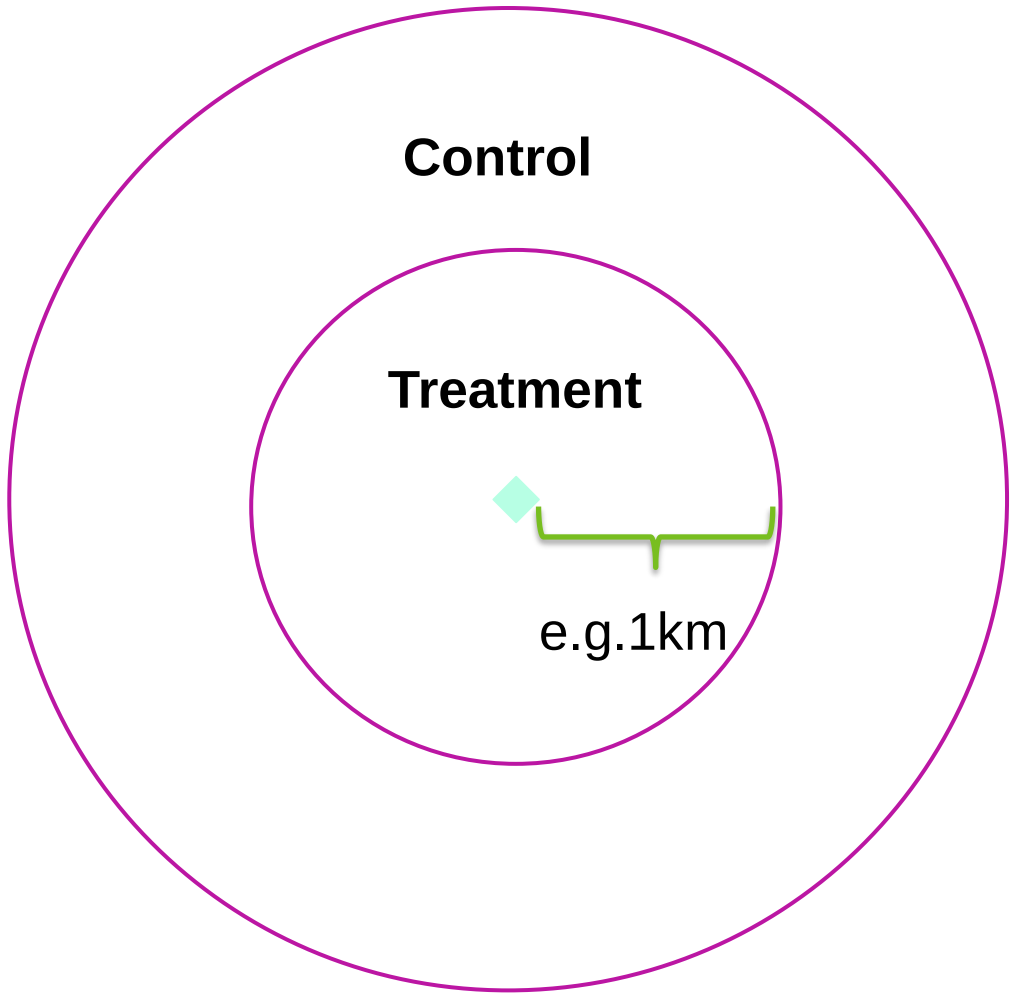
Fig. A3. A route map of the metro in Helsinki and Espoo after west metro is operational in 2017 (© HSL)

Harjunen (2018): West Metro extension and house prices



https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxvc2thcmloYXJqdW5lbnxneDoxY2JkNmZiMGM4ZWJmMjcx

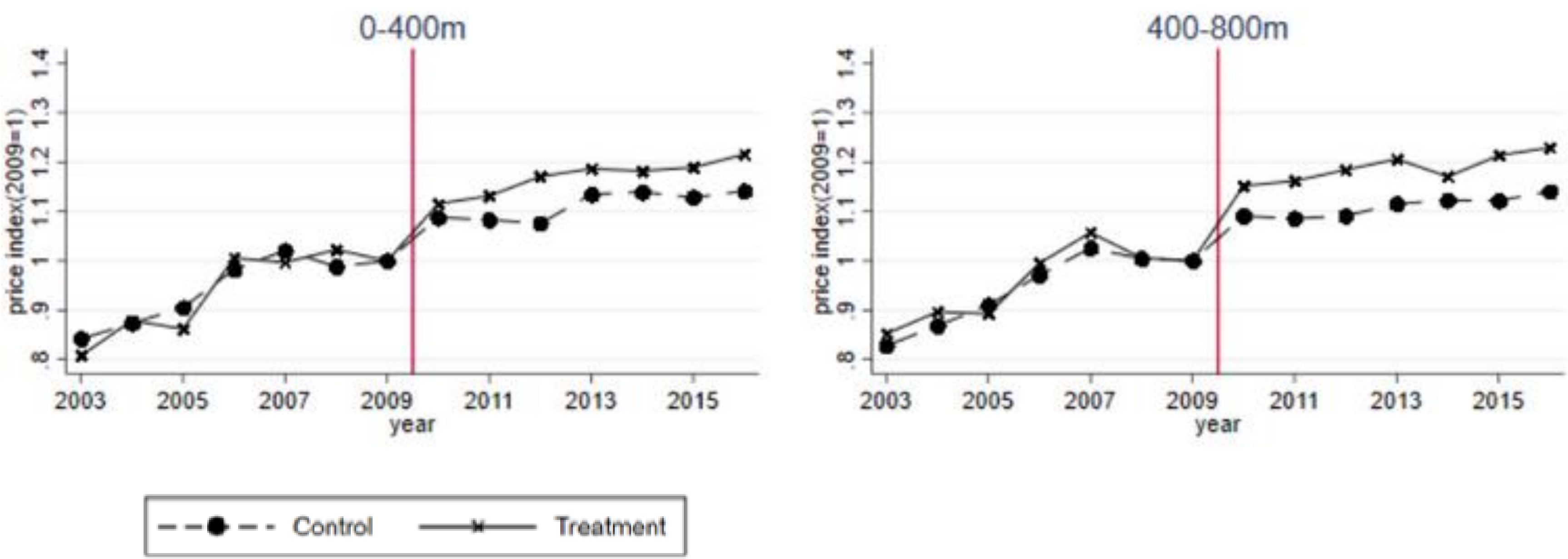
Is this a good DID design?



Data

Sample	Whole data (Helsinki and Espoo)	0 to 800m		800 to 1 600m	
Status		Treated	Control	Treated	Control
N	43 025	6 868	15 640	4 429	11 267
Sale price	223 668	252 024	196 154	311 661	199 122
	[110 007]	[119 458]	[78 980]	[156 343]	[82 107]
Square price	3 506	4 181	3 325	3 877	3 242
	[918]	[951]	[805]	[919]	[805]
Area	66	62	61	82	64
	[29]	[27]	[25]	[38]	[27]
Age	37	43	32	32	39
	[17]	[17]	[17]	[13]	[18]
Maint. Charge (€/m2)	3,5	3,8	3,5	3,5	3,5
	[1.2]	[1.1]	[1.2]	[1.2]	[1.3]
Floor number	2,4	2,7	2,5	2,3	2,3
	[1.6]	[1.7]	[1.5]	[1.5]	[1.4]
Floors in building	3,8	4,4	3,8	3,6	3,4
	[3.0]	[2.2]	[2.1]	[2.3]	[1.9]
Dist. to nearest station (m)	869	482	484	1 168	1 134
	[489]	[190]	[185]	[239]	[239]
Dist to CBD (km)	12	9	13	11,2	12,5
	[4.6]	[3.6]	[4.8]	[3.2]	[4.6]

Results – graphical

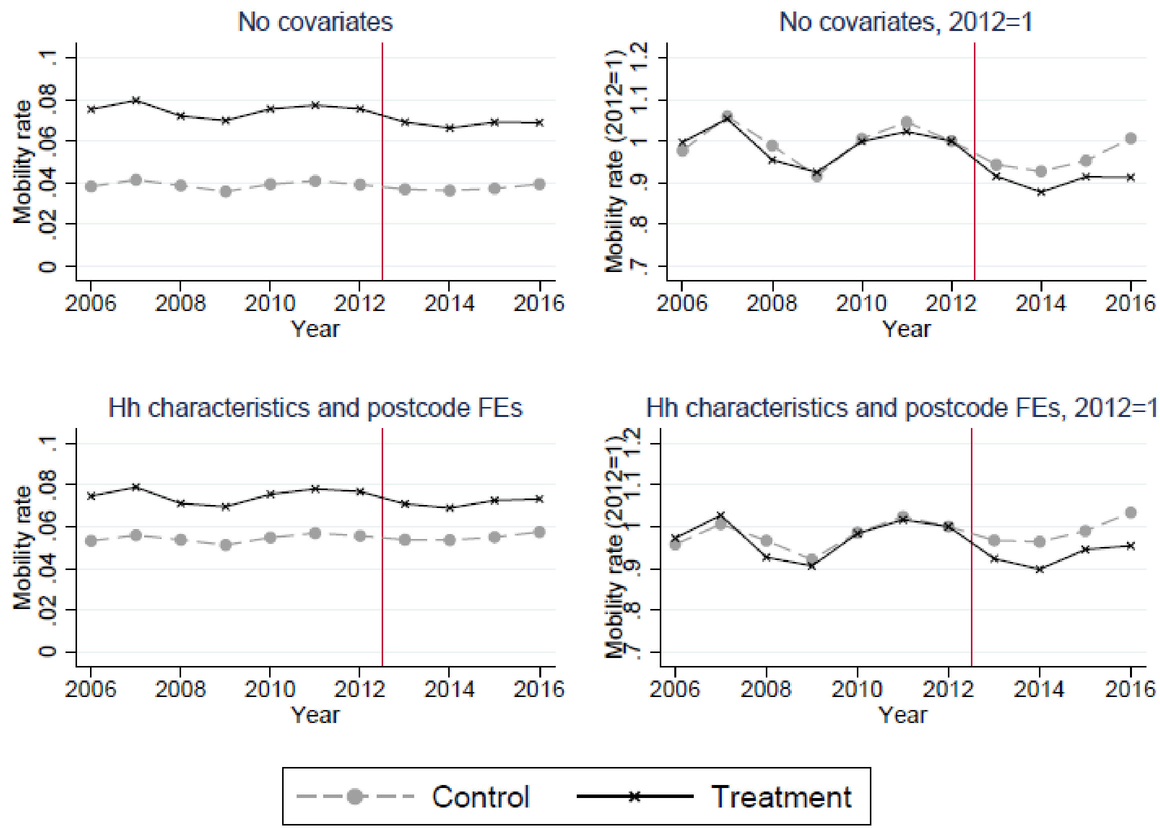


Eerola, Harjunen, Saarimaa & Lyytikäinen (2019)

- **Exploit a tax reform implemented in March 2013**
- **Raised the transfer tax rate of apartments in multi-unit buildings without affecting the tax rate of single-family detached houses => a DID design**
  - Treatment group = homeowners living in housing units subject to the tax increase (tax rate 1.6% -> 2%)
  - Control group = homeowners whose housing units were unaffected by the reform (tax rate constant at 4%)
  - Outcome: mobility, i.e. probability that the household moves
  - Data: all Finns 2006–2016

DID results

- **Is this a good design?**



https://www.cesifo.org/en/publikationen/2019/working-paper/effects-housing-transfer-taxes-household-mobility

What is the problem?

Table 2: Mobility rates before and after reform by origin and destination house type.

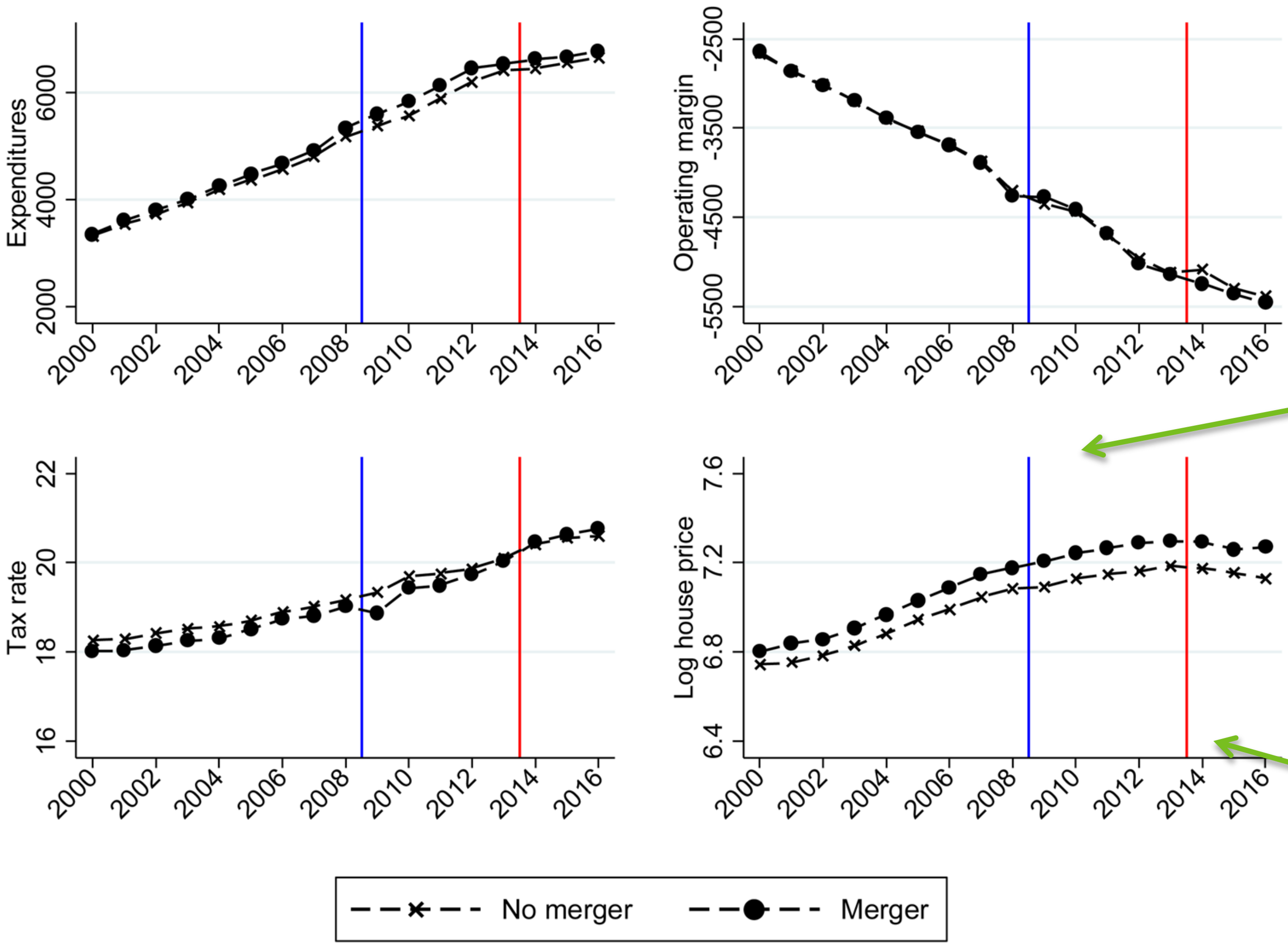
	House		Apartment	
	Pre-reform	Post-reform	Pre-reform	Post-reform
Moved to house	0.0144	0.0130	0.0193	0.0167
Moved to apartment	0.1012	0.0091	0.0334	0.0283
Moved to rental	0.1454	0.0153	0.0222	0.0233
Total	0.0391	0.0375	0.0749	0.0683

Notes: Table reports mobility rates of home-owners by origin and destination housing type.

Spillover to control group

- **In a housing market setting the design may be flawed due to spillover effects between the treatment and control groups**
  - If homeowners in the treatment group move less often because of the tax increase, the homeowners in the control group may also be indirectly affected as now they have less trading partners
- **Complement empirical analysis with a model with two homeowner segments, apartments and single-family houses**
  - Combining the empirical and theoretical analyses, we find a roughly 7.2% reduction in treatment group mobility due to a 0.5 percentage point increase in the transfer tax rate
  - Our DID estimate of the effect is roughly 5.6%, suggesting a 22% downward bias in the estimate. The bias arises because mobility decreases by 1.6% also in the control group.

Harjunen, Saarimaa & Tukiainen (2021)

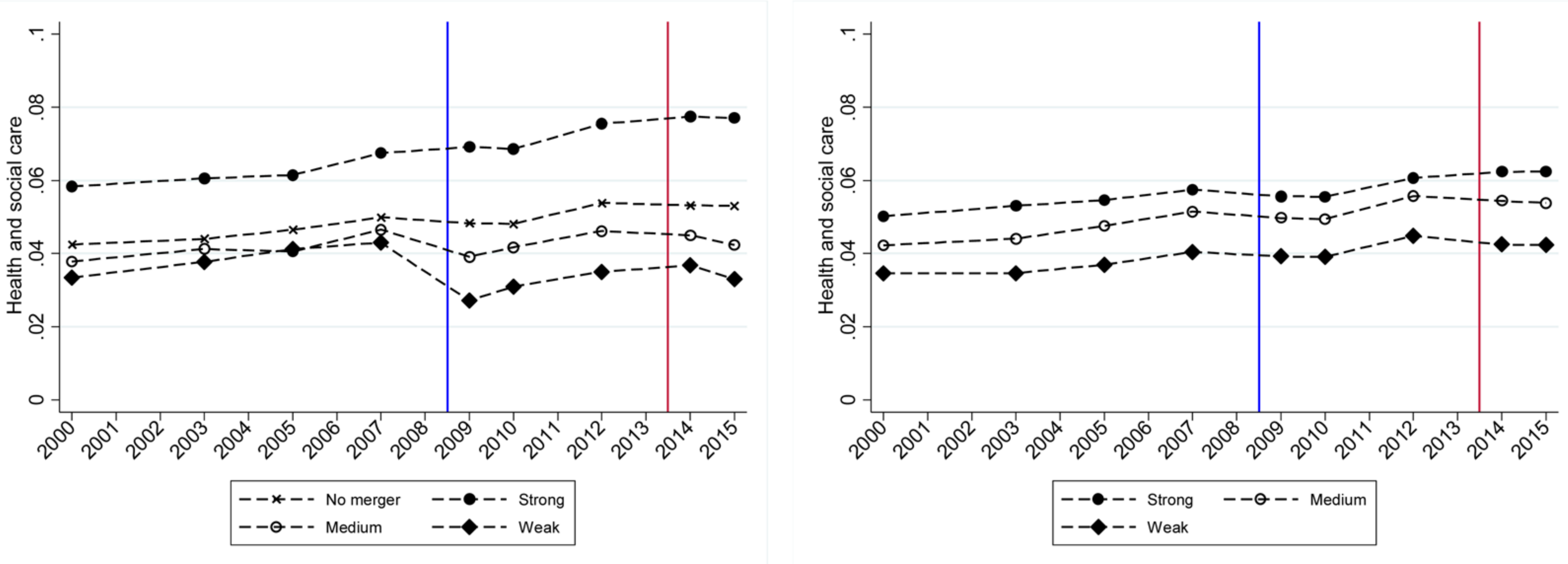


Mergers took place

A five year lay-off protection ended

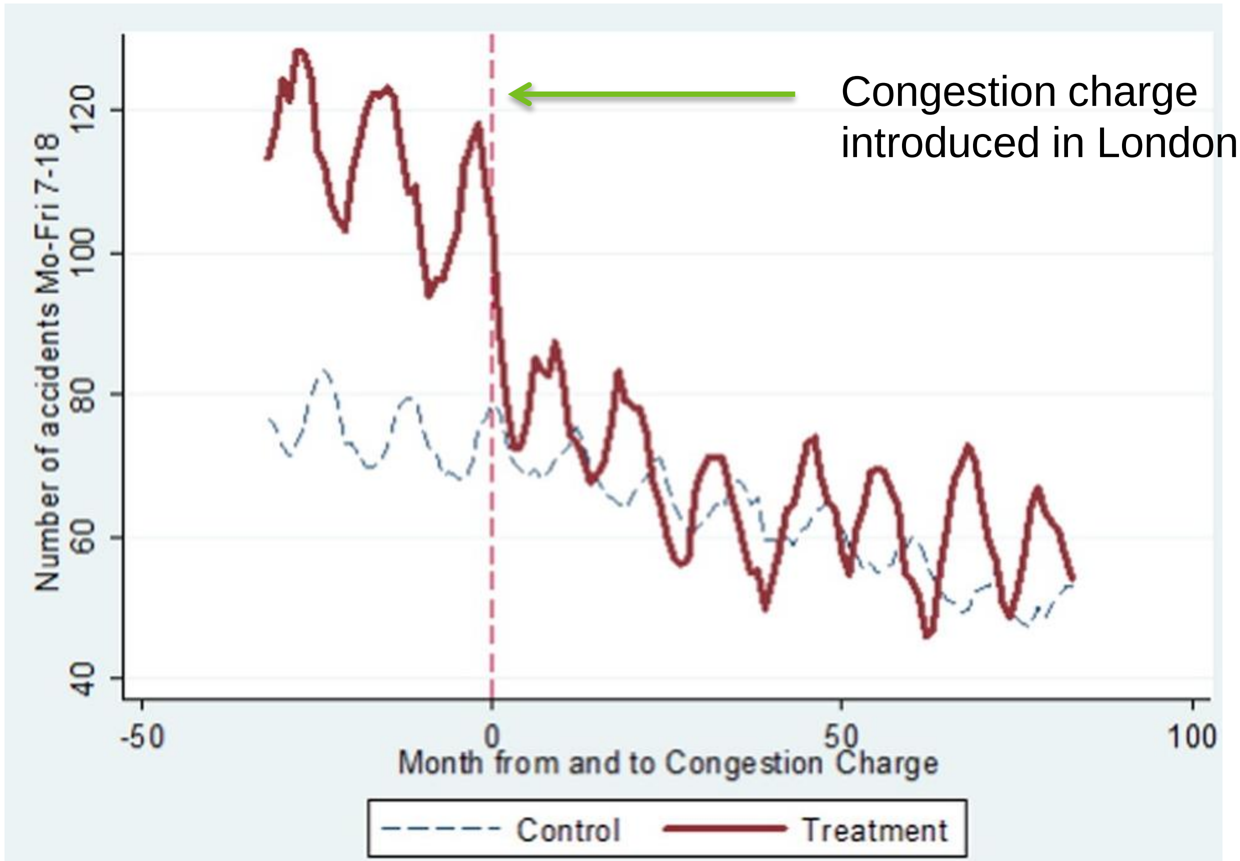
https://www.cambridge.org/core/journals/political-science-research-and-methods/article/political-representation-and-effects-of-municipal-mergers/1DC538037E1E3DC260EA276CD845318D#article

Main results – pre-merger municipality level



Notes: The left-hand side figure illustrate the true treatment effects based on municipal seat shares in the post-merger councils. The right-hand side figure illustrate the placebo treatments for the non-merged control group based on municipal population shares. The non-merged control group is constructed using nearest neighbor matching algorithm.

Green et al. (2016)

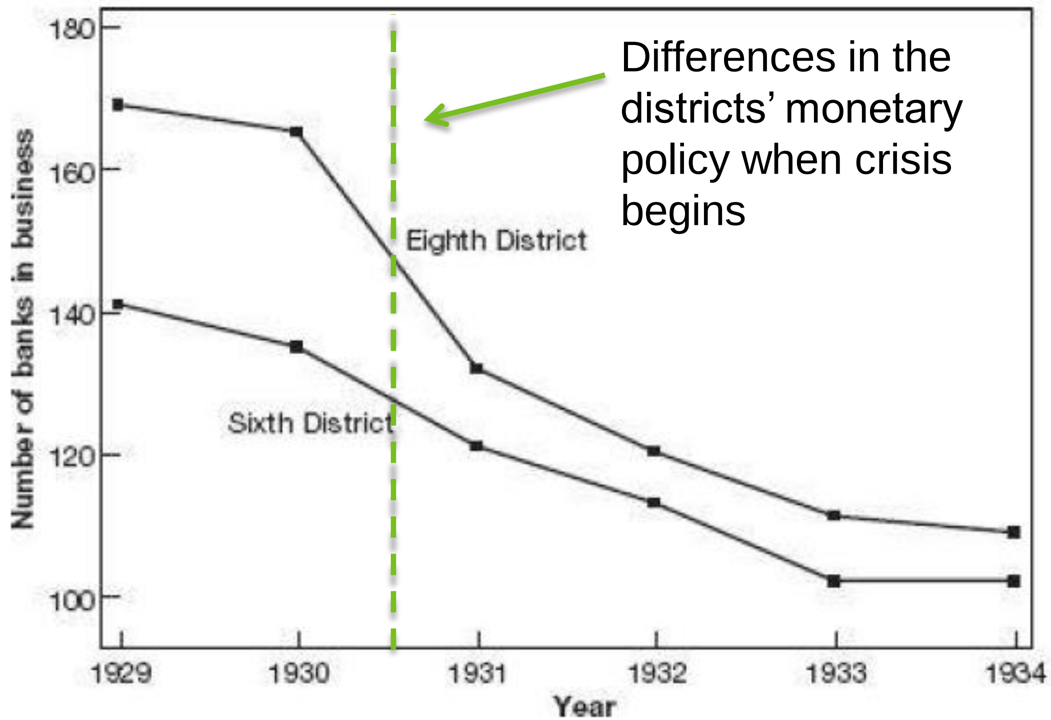


https://www.sciencedirect.com/science/article/pii/S0047272715001929

Richardson & Troots (2009)

FIGURE 5.2

Trends in bank failures in the Sixth and Eighth Federal Reserve Districts



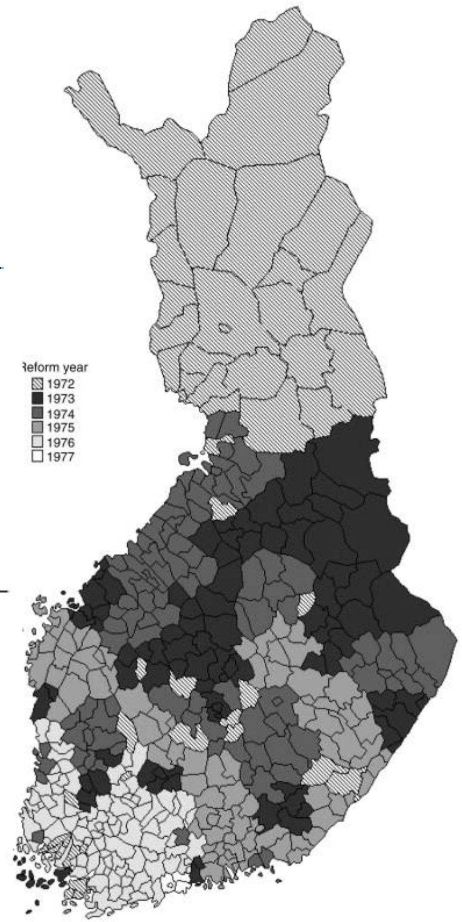
Note: This figure shows the number of banks in operation in Mississippi in the Sixth and Eighth Federal Reserve Districts between 1929 and 1934.

https://www.journals.uchicago.edu/doi/abs/10.1086/649603

Gradual implementation of a policy Pekkarinen et al. (2009)

Abstract

This paper estimates the effect of a major education reform on intergenerational income mobility. The Finnish comprehensive school reform of 1972–1977 replaced the old two-track school system with a uniform nine-year comprehensive school and shifted the selection of students to vocational and academic tracks from age 11 to age 16. We estimate the effect of this reform on the intergenerational income elasticity using a representative sample of males born between 1960 and 1966. The identification strategy relies on a differences-in-differences approach and exploits the fact that the reform was implemented gradually across the country during a six-year period. The results indicate that the reform reduced the intergenerational income elasticity by 23% from the pre-reform elasticity of 0.30 to post-reform elasticity of 0.23.



https://www.sciencedirect.com/science/article/pii/S0047272709000619

Gradual implementation of a policy
Böckerman et al. (2019)

Abstract

We exploit a large-scale natural experiment—the rollout of a nationwide electronic prescribing system in Finland—to study how digitization of prescriptions affects pharmaceutical use and health outcomes. We use comprehensive administrative data from patients treated with benzodiazepines, which are globally popular, effective but addictive psychotropic medications. We find no impact on benzodiazepine use on average, but among younger patients e-prescribing increases repeat prescription use. Younger patients' health outcomes do not improve but adverse outcomes, such as prescription drug abuse disorders and suicide attempts, increase dramatically. Improving access to medication through easier ordering may thus increase medication overuse.

Key words: health information technology, electronic prescribing, repeat prescriptions, inefficiency, medication overuse

JEL classes: H51, H75, I12, I18

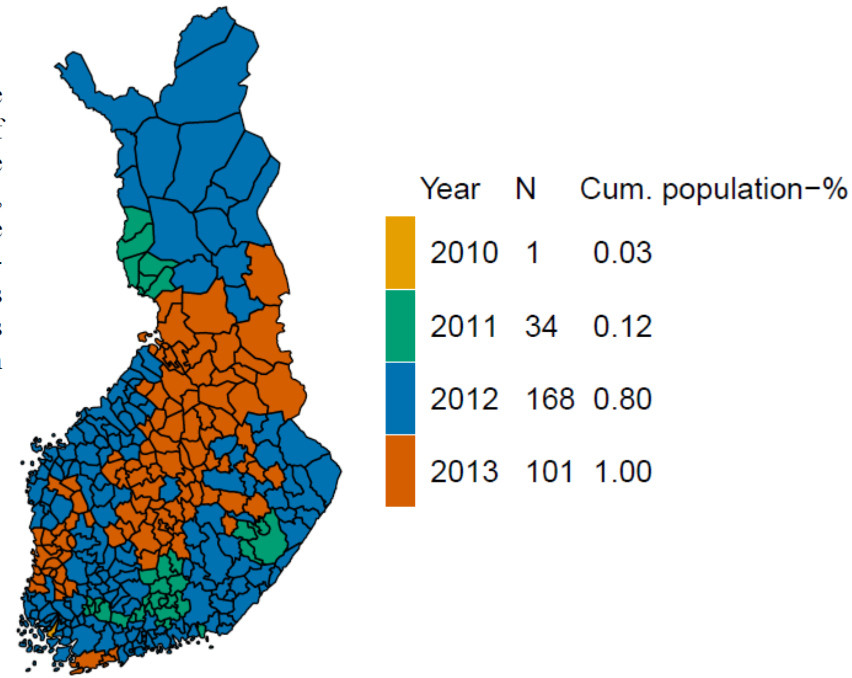
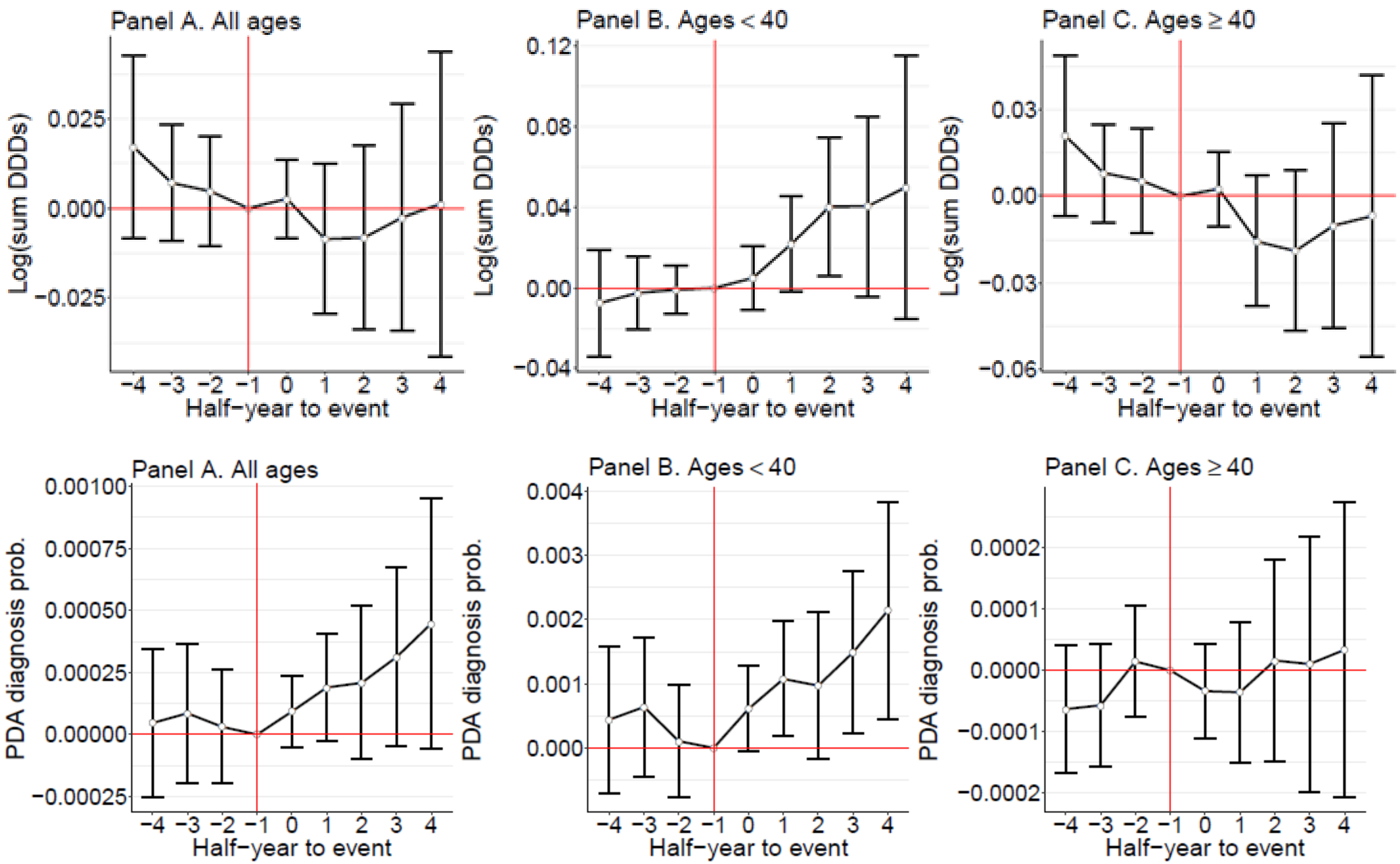


FIGURE 1: E-prescribing Adoption Year in Municipalities

https://vatt.fi/julkaisu?pubid=URN%3AISBN%3A978-952-274-235-3

Event study plots – Böckerman et al.



Further issues

Implementation

- **In practice, estimation of the treatment effect is implemented using regression models**
  - Produces "automatically" the estimate of the treatment effect and the standard error
- **Data can be either**
  - Panel data: data where you observe the same individuals (units) in multiple time periods
  - Repeated cross-sectional data: e.g. repeated random samples from a population where you observe different individuals in different time periods
- **There are complicated issues concerning implementation of event study type designs and the literature is moving forward on this (e.g. what is the control group at each stage?)**

DID recap

- **Idea:**
  - Even if treated and control groups differ in baseline characteristics, we can use observations on treatment and control groups before and after the treatment to estimate a causal effect
- **Assumptions:**
  - The potential outcomes (not observed) would have developed in a parallel manner for both groups in the absence of treatment
  - No spillovers and coinciding reforms
- **Testing for design validity:**
  - Visualization: are trends in outcomes parallel before treatment?
  - Is there anything else that could have happened to one group but not the other? (knowing your institutional setting)